



Alkalmazott nyelvtudomány 25. évf. 1. sz. (2025)

Tartalom

Tanulmányok

- VÁRADI Krisztián , PARAPATICS Andrea : A többnyelvűség tükröződése a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola nyelvi tájképében **1-23** [1.68 MB - PDF]
- JUHÁSZ Kornélia , DEME Andrea : A helyesírásban jelölt és nem jelölt j akusztikai megvalósulása gyors beszédtempóban **24-43** [678.96 kB - PDF]
- GÓSY Mária , VLADÁR Zsuzsa : Az izé eredete, története és mai használata **44-66** [543.49 kB - PDF]
- DODÉ Réka , YANG Zijian Győző : Generálás vs. kinyerés: nagy nyelvi modellel kapott kulcsszavak vizsgálata **67-83** [374.24 kB - PDF]
- Adrianna KOSTIC : Minorities in the digital space: Translanguaging practices of Vojvodina Hungarians online **84-103** en [4.43 MB - PDF]
- Wakwoya MIHIRETU , Margaret DEUCHAR : Analyzing Code-Switching Between Afaan Oromoo and English in Ethiopia: A Grammatical Perspective **104-121** en [403.32 kB - PDF]
- Anita KISS : Contrastive analysis of attitudes towards dialects **122-136** en [398.57 kB - PDF]
- Csilla Ibolya SÓLYOM : Transfer of pronominal possessors, differential object marking, and resultative constructions in L3 acquisition of English **137-160** en [462.63 kB - PDF]

- **Annamária KISS , Valéria CSÉPE : Investigating working memory impairments in bilingual aphasia 161-186 en [1.10 MB - PDF]**
 - **Merga Feyera WEKJIRA , Anat STAVANS : Profiles of multilingual agencies in educational contexts in Oromia, Ethiopia 187-209 en [494.64 kB - PDF]**
-

Könyvszemle

- **JAKAB Dorottya : XIX. Alkalmazott Nyelvészeti Doktoranduszkonferencia 210-212 [181.66 kB - PDF]**
- **ANDOR József : Ara Rauch (szerk.): Szókereső írószótár. Magyar főnevek szókapcsolatai 213-216 [196.97 kB - PDF]**
- **BALÁZS Boróka : Svindt Veronika - Bóna Judit (szerk.): Általános nyelvészeti tanulmányok XXXVI. Tanulmányok az anyanyelv-elsajátítás köréből 217-219 [168.10 kB - PDF]**
- **LŐRINCZ Julianna : Domonkosi Ágnes: A megszólítás gyakorlatai. Tegezés, magázás, szólítás a társas viszonyok alakításában 220-223 [175.99 kB - PDF]**
- **LUTA Adrienn : Terminológiai csütörtök. XI. Szakmai nap: Terminológia, fordítás, szakszövegírás. Lehetőségek és elvárások a munkaerőpiacon 224-226 [165.39 kB - PDF]**
- **ULLMANN Cintia : Tamás Dóra Mária - Szoták Szilvia (szerk.): Terminológiastratégia és jogi alapterminusok a szomszédos országok nyelvén 227-230 [171.85 kB - PDF]**
- **SURÁNYI Anna : Karen Stollznow: Bitch: The Journey of a Word 231-234 [175.30 kB - PDF]**
- **Nilay Nur TAŞDEMİR : Preserving the past, shaping the future: the sixth international workshop on the history of speech communication research (HSCR 2024) 235-238 en [176.56 kB - PDF]**
- **Adrienn BIRÓ : Rebecca Tipton: The Routledge Guide to Teaching Ethics in Translation and**

VÁRADI KRISZTIÁN¹ – PARAPATICS ANDREA²

¹Pannon Egyetem, Többnyelvűségi Nyelvtudományi Doktori Iskola; II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola, Filológia Tanszék
varadi.krisztian@kmf.org.ua
<https://orcid.org/0000-0001-7554-4158>

²Pannon Egyetem, Magyar és Alkalmazott Nyelvtudományi Intézet
parapatics.andrea@htk.uni-pannon.hu
<https://orcid.org/0000-0002-1711-0885>

Váradi Krisztián – Parapatics Andrea: A többnyelvűség tükröződése a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola nyelvi tájképében
Alkalmazott Nyelvtudomány, XXV. évfolyam, 2025/1. szám, 1–23.
doi:<http://dx.doi.org/10.18460/ANY.2025.1.001>

A többnyelvűség tükröződése a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola nyelvi tájképében

A distinct type of linguistic landscape study is the so-called schoolscape analysis, which examines how different languages are divided and used in the visual space of educational institutions. The present study aimed to analyse the presence of bi- and multilingualism in the schoolscape of the Ferenc Rákóczi II Transcarpathian Hungarian College of Higher Education. The institution is a private higher educational institution in Ukraine that uses Hungarian as its primary language of instruction, thus providing a fertile ground to investigate how Hungarian, as a minority language, Ukrainian, as a state language, and English as an international language, are divided into the formal and informal domains of the college.

The research employed a longitudinal approach, documenting linguistic signs over ten months from March to December 2024. As a result, a corpus consisting of 546 pictures was compiled. Additionally, a virtual schoolscape analysis of six websites and social networking sites associated with the college was also included. The collected pictures were divided into nine categories ranging from bottom-up (informal) to top-down (formal) signs (Ben-Rafael et al., 2006). These categories included decorations, symbolic inscriptions and signs, advertisements, library signage, instructions related to crises, room names, educational documents, institutional nameplates, and other linguistic signs.

Our results showed that the analysed college follows a strong bilingual tradition by making all important information available in both Hungarian and Ukrainian, including timetables, lists of thesis topics, semester structures, escape plans, room names, and most institutional advertisements. The college even has a bilingual learning management system, called IRIS, and shares news on its website and Facebook page in three languages, including Hungarian, Ukrainian, and English. However, a difference in the number of monolingual Hungarian and Ukrainian signs was observed between the bottom-up and top-down domains. When students or college employees produced their own signs, they mostly preferred to use the Hungarian language alone. For instance, almost all the graduation photo collages and some informative signs, such as the office hours of the library during the summer, were in Hungarian monolingual. The same can be said about the menus in the canteens, but interestingly, when they shared the daily menu on their Facebook pages, it was bilingual in Hungarian and Ukrainian.

There were also cases where, although the Hungarian language was more dominant, the Ukrainian state language was present, too. For example, the labels on the bookshelves in the library were more detailed in Hungarian by denoting all the scientific subfields to which the books belonged while indicating only the main fields in Ukrainian beneath them. Advertisements also showed a significant amount of variety because most of them were Hungarian–Ukrainian bilingual, either by showcasing information in two languages on the same sign or by making two identical signs in two different

languages, but many advertisements were only available in Hungarian, including the posters created for the admission campaign of the college (although it was available online in Ukrainian too). For these reasons, the informal bottom-up sphere of the college was more Hungarian-dominant.

When discussing top-down signs, nearly all of them were available in both Ukrainian and Hungarian, with the Ukrainian version typically displayed above or to the left of the Hungarian text, in accordance with Ukraine's language laws. For instance, signs denoting procedures in case of crises, institutional nameplates, and legal documents were all predominantly in Ukrainian. Additionally, Ukrainian national symbols were much more frequently observed in the college.

In conclusion, the Ferenc Rákóczi II Transcarpathian Hungarian College of Higher Education is a predominantly bilingual institution that emphasises the acceptance of different languages and cultures through Hungarian–Ukrainian artwork and the translation of nearly everything into both languages. The Hungarian language is more dominant in the bottom-up sphere, while Ukrainian–Hungarian bilingualism is the most frequent in the top-down domain, and the use of the English language is minimal. Future research may yield more interesting results about the symbolic functions of different languages in other educational institutions, as this case study provides only a glimpse into the complex linguistic situation in Transcarpathia, Ukraine.

Keywords: Hungarian minority, linguistic landscape, multilingualism, schoolscape, Transcarpathia

1. Bevezetés

Kárpátalja egy többnyelvű régió Ukrajna nyugati felében, ahol a legutóbbi hivatalos népszámlálási adatok (2001) alapján több mint száz nemzetiség él egymás mellett. A legnagyobb kisebbségi csoportot a magyarok alkotják, akik 151 516 fővel a lakosság 12,1%-át tették ki az évezred elején (Molnár & Molnár D., 2005). Ezt az adatot érdemes fenntartásokkal kezelni, mivel a gazdasági hanyatlás és a 2022. február 24-én kitört orosz–ukrán háború következtében a magyar népesség nagy része külföldre emigrált (Csernicskó & Gazdag, 2023).

Kárpátalja mindig is sokszínű régiónak számított mind nyelvi, mind kulturális szempontból. Ennek egyrészt földrajzi, másrészt történelmi okai vannak, hiszen a terület az elmúlt évszázadban négy különböző országhoz tartozott: Magyar Királyság, Csehszlovák Köztársaság, Szovjetunió, 1991-től Ukrajna (Csernicskó et al., 2023). A folyamatos államváltozások mellett a régióban hivatalos státusszal rendelkező nyelvek is rendszeresen váltakoztak. Napjainkban az ukrán az államnyelv, és a szovjet időszakban hivatalosnak számító orosz nyelv nyilvános használata be van tiltva. A magyar nyelv helyzete Ukrajnában egyre romlik az elmúlt évek megszűkítő intézkedéseinek következtében.

A kárpátaljai lakosság többsége két- vagy többnyelvű, és ez a nyelvi változatosság tükröződik az élet minden területén: vannak magyar tannyelvű iskolák, az emberek két vagy több nyelven kommunikálnak, valamint a magyarul lakta települések utcáin gyakoriak a kétnyelvű kiírások. Számos tanulmányban vizsgálták már Kárpátalja nyelvi tájképét, különösen az utolsó nagyszámú magyar lakossággal bíró városban, Beregszászon és annak vonzáskörzetében (Csernicskó, 2015, 2016; Hires-László, 2019; Hires-László et al., 2023; Tóth-Orosz, 2020).

Ezen belül az iskolai nyelvtájkép-vizsgálatok viszonylag új kutatási területnek számítanak Kárpátalján, bár néhány példát találhatunk a szakirodalomban

(Karmacsi, 2018; Karmacsi & Mészár, 2024; Ohár, 2022; Váradi et al., 2024). Mivel a kutatások többsége magyar tannyelvű középiskolákra és gimnáziumokra korlátozódik, a jelen tanulmányban egy magán felsőoktatási intézmény, a Beregszász városában található II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola nyelvi tájképét mutatjuk be. A Rákóczi-főiskola 1993-ban jött létre, de önálló oktatási intézményként csak 1996 óta működik. A 2024/2025. tanév 1386 hallgatóval indult. Mivel az intézmény többnyelvű környezetben helyezkedik el, a hallgatók többsége is legalább két nyelven kommunikál, de sokan akár három vagy több nyelven is, beleértve a magyar, ukrán, angol, német és orosz nyelveket (Váradi, 2023).

2. A nyelvitájkép-elemzések elméleti háttere

Nyelvi tájkép alatt különböző nyelvek előfordulását és megoszlását értjük egy terület nyilvános jelzésein, tábláin (Landry & Bourhis, 1997). A nyelvitájkép-elemeknek két alapvető funkciója van. A kommunikatív funkció értelmében a feliratok nyelve szimplán az információátadás eszköze. A szimbolikus funkció ennél jóval összetettebb, hiszen a nyilvános jelzések tükrözhetik az adott területen használt nyelvek státuszát és funkcióit (Cenoz & Gorter, 2006).

A nyelvi tájkép elemzése fényt deríthet a különböző nyelvek közötti hatalmi viszonyokra, különösképpen az állam- és a kisebbségi nyelvek megjelenésére a nyilvános térben (Barni & Bagna, 2015). A többnyelvű régiók nyelvi tájképének vizsgálata révén különböző társadalmi és politikai jelenségekre is magyarázatot találhatunk. Például egy kisebbségi tannyelvű iskola nyelvitájkép-elemzésével jobban megismerhetjük egy ország nyelvpolitikáját (Hires-László, 2019).

A nyelvitájkép-kutatások két megközelítés szerint végezhetők. Kvantitatív elemzés során az összegyűjtött nyelvi jeleket különböző szempontok szerint kategorizálják, így vizsgálva a nyelvek statisztikai megoszlását és előfordulási gyakoriságát. Ezzel szemben a kvalitatív nyelvitájkép-elemzések a jelzéseken megfigyelhető nyelvek elhelyezkedésének és gyakoriságának szimbolikus jelentőségét vizsgálják (Hires-László, 2019; Shohamy & Ben-Rafael, 2015).

A formalitásuk mértéke alapján megkülönböztetünk alulról felfelé (*bottom-up*) és fentről lefelé (*top-down*) irányuló szabályozás útján létrejövő nyelvitájkép-elemeket (Ben-Rafael et al., 2006). Az előbbi kategóriába tartoznak többek között a magánszemélyek és magánvállalkozások által önkényesen elhelyezett táblák és nyelvi jelzések, amelyek sokkal inkább informális jellegűek, mint az utóbbi csoportba tartozó kiírások, amelyek elhelyezését és nyelvezetét a „felülről jövő” állami és önkormányzati jogszabályok határozzák meg.

3. Iskolai nyelvitájkép-kutatások

Iskolai kontextusban a nyelvi tájképet úgy értelmezik, mint az oktatási intézmények fizikai környezetének nyelvészeti szempontú vizsgálatát, mivel az ott fellelhető szövegek és szimbólumok különböző nyelvi ideológiákat közvetítenek a tanulók felé (Brown, 2012). Az iskolai nyelvi tájképre a nemzetközi szakirodalomban *schoolscape*-ként hivatkoznak, amely a tanítás és tanulás fizikai és társas közegét jelenti (Brown, 2005). Különösen olyan intézményekben szolgálhat hasznos eredményekkel az iskolai nyelvitájkép-kutatás, ahol az államnyelv és különböző idegen nyelvek kíséretében kisebbségi nyelvek is jelen vannak (Bartha et al., 2013).

Az oktatási intézmények fontos szerepet játszanak a tudás átadásában, a kultúra megőrzésében és a társas kapcsolatok kifejlesztésében, ezért is kiemelten fontos az iskolák falain belül megfigyelhető nyelvi jelek vizsgálata (Wedin, 2021). Egy Baszkföldön végzett iskolai nyelvitájkép-vizsgálat (Gorter & Cenoz, 2015) során a kutatók a nyelvi jeleket kilenc kategóriába csoportosították az informatív és szimbolikus funkcióik alapján. Következtetéseik szerint a többnyelvű kiírások és jelzések hozzájárulnak a diákok többnyelvű kompetenciáinak a fejlesztéséhez. Az iskolában eltöltött évek kétségkívül fontos szerepet játszanak a fiatalok nyelvi szocializációjában.

Iskolai nyelvitájkép-elemzések készültek már Magyarországon (Szabó, 2015), valamint magyar tannyelvű romániai (Biró, 2016; Laihonon & Tódor, 2017; Tódor, 2019) és szlovákiai (Lőrincz, 2021) oktatási intézményekben is. Kárpátalján elsősorban a Beregszászi kistérség magyar tanítási nyelvű középiskoláira és egyházi fenntartású líceumaira fókuszálnak a kutatók (Karmacsi, 2018; Karmacsi & Mészár, 2024; Ohár, 2022). A magyar kisebbség felsőoktatási intézményeinek iskolai nyelvi tájképét a legjobb tudomásunk szerint csak egy tanulmány keretében vizsgálták eddig (Váradi et al., 2024), így írásunkkal ezt a kutatási területet szeretnénk bővíteni.

A vizsgálatok egy új iránya az oktatási intézmények weboldalain megjelenő információk vizsgálata, vagyis a virtuális iskolai nyelvitájkép-elemzés (Lőrincz et al., 2021; Veres, 2024). Ez azért is fontos, mert egy oktatási intézmény nyelvi tájképe különböző szintekből épül fel attól függően, hogy az adott szinten megjelenő nyelvi jelek mekkora nyilvánosságot érnek el. A funkcionális színteződés modellje (Balog, 2023; Istók et al., 2024) az iskolai nyelvi tájképet négy szintre osztja: 1) mikrostruktúra (osztálytermek); 2) mezostruktúra (folyosók); 3) makrostruktúra (az iskola külső falai); 4) megastruktúra (az iskola weboldala). Érdekes szem előtt tartani, hogy nem minden oktatási intézménynek van weboldala, így néhány esetben csak három funkcionális szintről beszélhetünk. A hazai, határon túli és nemzetközi iskolai nyelvitájkép-kutatásokról Váradi Krisztián (2025) tanulmányában lehet részletesebben olvasni.

4. Jogszabályi háttér

Egy amerikai iskolában folytatott kutatás során Johnson (1980) arra jutott, hogy a falakon fellelhető szimbólumok a nacionalizmust és az amerikai öntudatot erősítik a tanulóknál. Ukrajnában az utóbbi másfél évtizedben erős ukránosítási politika zajlik, így a nemzeti kisebbségek iskoláinak nyelvitájkép-elemzése kiemelt jelentőséggel bír. A jogszabályok értelmében az országban minden állami és önkormányzati intézmény névtáblája elsődlegesen az államnyelven, ukránul készül el, azonban azokon a területeken, ahol egy adott nemzeti kisebbség aránya eléri a 10%-ot a lakosságon belül, megengedett a kisebbségi nyelv használata is az intézményi névtáblákon az államnyelvi szöveg alatt vagy mellett, kisebb betűmérettel (Karmacsi, 2018).

A helyzet az utóbbi években sajnos sokat romlott amiatt, hogy az évtizedekig fennálló helyi közigazgatási egységeket kistérségekbe szervezték át. Ennek következtében számos magyartöbbségű település került olyan adminisztratív egységbe, amelynek össznépességében a magyar kisebbség aránya nem éri el a 10%-ot (Tóth, 2024). A 2017-es oktatási törvény¹ és a 2019-es államnyelvi törvény² jelentősen korlátozta a magyar tannyelvű oktatás jelenét és jövőjét, hiszen előírták az ukrán államnyelv egyre magasabb százalékban történő alkalmazását az oktatási folyamatban. Még a magán felsőoktatási intézmények szabad tannyelvválasztási jogát is elvették, sőt az önkormányzati hivatalok kétnyelvű névtábláit is elkezdtek lecserélni ukrán egynyelvű táblákra. Ezen túl a korábban magyar–ukrán kétnyelvű helynévtáblákon az ukrán államnyelv mellett csak a cirill betűs megnevezések latin betűs átíratai jelenhetnek meg (Csernicskó, 2022).

Az Európai Unióhoz való csatlakozás biztosítása érdekében 2023 decemberében némileg enyhítettek az Európai Unió valamelyik hivatalos nyelvét beszélő nemzeti kisebbségek oktatási és nyelvi jogainak korlátozásain. Így államnyelven már csak Ukrajna történelmét és a honvédelmet, valamint értelemszerűen az ukrán nyelvet és irodalmat kell tanulniuk a kisebbségi diákoknak (Fedinec, 2024). A felsőoktatásról szóló törvény³ 48. cikke 4. pontjának értelmében a magán felsőoktatási intézmények – mint a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola is – visszakapták a tannyelvük szabad megválasztásának lehetőségét azzal a feltétellel, hogy kötelező jelleggel biztosítják az ukrán államnyelv oktatását.

¹ <https://zakon.rada.gov.ua/laws/show/2145-19>

² <https://zakon.rada.gov.ua/laws/show/2704-19>

³ <https://zakon.rada.gov.ua/laws/show/1556-18>

5. Kutatási módszer

A jelen tanulmány célja, hogy bemutassa a két- és többnyelvűség tükröződését egy kárpátaljai magyar felsőoktatási intézmény, a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola nyelvi tájképében. Különös figyelmet fordítunk a magyar mint kisebbségi nyelv, az ukrán mint államnyelv és az angol mint idegen nyelv közötti viszony vizsgálatára, kitérve azok szimbolikus funkciójára és jelentőségére egy alapvetően kisebbségi tannyelvű felsőoktatási intézményben.

Az adatgyűjtés 2024 márciusa és decembere között zajlott tíz hónapon keresztül. A longitudinális megközelítésre azért volt szükség, hogy nyomon tudjuk követni a nyelvi tájképben bekövetkező változásokat, ha vannak, ezáltal biztosítva a kutatási eredmények megbízhatóságát. A tanulmány első szerzője minden nyelvi jelről fényképeket készített az intézmény főépületében, a hozzá tartozó Filológia Campuson, valamint a főiskola Egészségügyi és Sportcentrumában. Az így összeállított korpusz 546 fényképet tartalmaz, amelybe beletartozik a funkcionális szinteződés modellje alapján az iskolai nyelvi tájkép mikro-, mezo-, makro- és megastruktúrája is (lásd fent). Adatolt nyelvitájkép-elemnek számítunk minden üzenettel bíró vizuális elemet, beleértve a feliratokat, képeket és szimbólumokat. Emellett figyelmet fordítottunk a Rákóczi-főiskola weboldalának és közösségi média oldalainak (összesen 6 internetes oldal) virtuális nyelvitájkép-elemzésére is (Biró, 2018). Erre azért volt szükség, mert egy oktatási intézmény weboldala bír a legnagyobb közönséggel az összes nyelvitájkép-elem közül. Mindezekkel együtt az elemzésre szánt nyelvitájkép-elemek száma összesen 552 lett. Mivel célunk elsősorban a különböző nyelvek szimbolikus jelentőségének a vizsgálata volt, a kvalitatív kutatási megközelítést alkalmaztuk.

6. A Rákóczi-főiskola nyelvi tájképe

A fényképeket kiválogattuk és kilenc kategóriába soroltuk a rajtuk szereplő nyelvi jelzések funkciója szerint, amihez tizedik kategóriaként a virtuális nyelvi tájkép csatlakozott (1. táblázat).

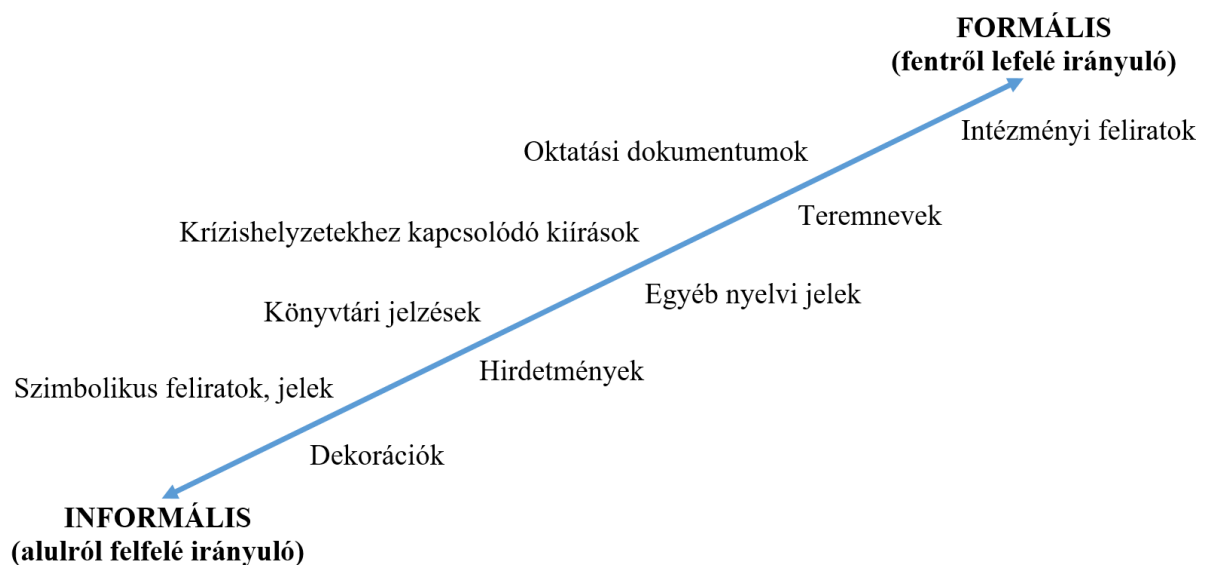
1. táblázat. Az iskolai nyelvitájkép-elemek csoportosítása (Saját szerkesztés)

Kategóriák	Összetevők	Mennyiség
Dekorációk	Műalkotások, festmények, informatív posztterek, tablók, időszakos kiállítások	173
Szimbolikus feliratok, jelek	Szobrok, emléktáblák, nemzeti jelképek (zászló, címer, himnusz)	16
Hirdetmények	Belső intézményi hirdetmények, külső álláshirdetések	98
Könyvtári jelzések	Könyvtári feliratok, könyvespolcok elrendezése, tudományos adatbázisokhoz kapcsolódó plakátok	36

Kategóriák	Összetevők	Mennyiség
Krízishelyzetekhez kapcsolódó kiírások	A COVID-19 járványhoz, az orosz–ukrán háborúhoz és más vészhelyzetekhez kötődő jelzések	38
Teremnevek	Az osztálytermek, irodák, tanszékek névtáblái	43
Oktatási dokumentumok	Témakiírások, tantermen kívüli foglalkozások listája, órarendek, félévszerkezetek, tanszéki dokumentáció	58
Intézményi feliratok	Intézményi névtáblák, alegységek külső névtáblái	7
Egyéb nyelvi jelek	Jogi dokumentumok, kijáratokat és mosdókat jelző táblák, menük, egyéb informatív kiírások	77
Virtuális iskolai nyelvi tájkép	A főiskola weboldala, publikációtára, közösségi média oldalai, a Kiscsikó menza Facebook-oldala	6
Összesen		552

A nyelvi jelzések az alulról felfelé szabályozott (informális) jelektől a fentről lefelé irányuló (formális) nyelvitájkép-elemekig terjednek (Ben-Rafael et al., 2006), ezért a funkciókat az 1. ábrán szemléltetve is felosztottuk. Ez a felosztás remélhetőleg több iskolai nyelvitájkép-kutató számára is hasznosítható lesz a jövőben. Az egyéb nyelvi jelek azért kerültek az ábra közepére, mert egyaránt tartoznak ide alulról felfelé és felülről lefelé szabályozott jelzések.

1. ábra. Az iskolai nyelvitájkép-elemek formalitás szerinti csoportosítása
(Saját szerkesztés, Ben-Rafael et al., 2006 alapján)



A következőkben külön alpontokban elemezzük az így kategorizált adatokat. Valamennyi szemléltetésül beszúrt ábra és a rajtuk szereplő kép forrása a tanulmány első szerzője, aki a vizsgált intézmény egyik oktatója is egyben. A tanulmány második szerzője az adatok elemzésében nyújtott segítséget és adott hasznos tanácsokat.

6.1. Dekorációk

A dekorációk kétségkívül az iskolai nyelvi tájkép informális részét képezik. Ebbe a kategóriába soroltunk minden olyan elemet a folyosók és tantermek falairól, amelyek pusztán díszítő funkciót töltek be, és szerepelt rajtuk valamilyen felirat. Például: informatív poszterek, műalkotások, festmények, térképek, tablók, közetgyűjtemények, időszakos kiállítások.

Biró (2016) megközelítése szerint a nyelvitájkép-elemek élettartamuk szerint lehetnek állandóak (pl. nemzeti zászló), hosszú távúak (pl. tanév rendje) és rövid távúak (pl. hirdetmények). A dekorációk az állandó jelzések közé tartoznak, mivel akár több éven keresztül sem cserélik le őket.

Csak magyar nyelvű feliratokat tartalmaztak például a híres magyar feltalálókról készített poszterek a folyosókon, az Aranycsapat tagjait bemutató fényképsorozat feliratai a főiskola Sportcentrumában, illetve a Magyar Tanszék osztálytermeiben a magyar nyelv grammatikai szabályait összefoglaló szemléltetők. Ezek mind a magyar kultúra valamely szegmensét közvetítik a hallgatók felé. A legtöbb tablón magyar nyelven szerepelt a végzett hallgatók neve és az általuk választott idézet, még az Ukrán Tanszék végzőseinek tablóján is. Egyedül az Angol Tanszék tablóján fordult elő angol nyelvű idézet.

A magyar–ukrán kétnyelvű feliratok szintén fontos szerephez jutottak a falakon található dekorációknál. A feliratok helyzete alapján a magyar nyelv dominanciája figyelhető meg, mivel egymás melletti elhelyezkedés esetén a magyar változat szerepelt bal oldalon, egymás alatti elhelyezkedéskor pedig a magyar szöveg volt felül (2. ábra). A legtöbb kétnyelvű dekorációnak specifikusan kárpátaljai vonatkozása volt, például kárpátaljai települések pecsétjei; kárpátaljai művészek festményei, alattuk kétnyelvű feliratokkal; a Magyar Királyság egykori megyéinek címerei.

2. ábra. Kétnyelvű dekorációk



Külön figyelmet érdemel a Rákóczi-főiskola Művészeti Részlegén található kétnyelvű műalkotás, amely magyar és ukrán nyelven is tartalmazza a részleg nevét, a főiskolához kötődő kárpátaljai művészek életrajzát, az egyik festett emberalak közepén pedig egy szív látható, amelyben két nyelven szerepel az „Itt a helyed” felirat. A falfestmény további része egy női alak, aki egy vizeskancsót tart a kezében, amelyből a kiömlő vízben megtalálható kék betűkkel a víz szó különböző nyelveken, ezzel is a többnyelvűsége és a multikulturalizmusra helyezve a hangsúlyt (3. ábra). Ezek a dekoratív elemek mind azt hivatottak jelképezni, hogy az intézményben anyanyelvtől függetlenül mindenki otthon érezheti magát.

3. ábra. Kétnyelvű művészeti alkotások



A kompozíció folytatásaként egy könyvespolc van festve a szomszédos falra, amelyen megtalálhatók az ukrán, a magyar, a kárpátaljai és a világirodalom nagyjainak a művei is. A festett könyvespolc mellett egy valódi, vitrinnel ellátott polc is van, amelyen bárki által kölcsönvehető magyar, ukrán és angol nyelvű könyvek sorakoznak. A kölcsönzésnek csak egy feltétele van: a könyv elolvasása után a három nyelv valamelyikén ki kell tölteni egy kis könyvjelzőt, így ajánlva az adott olvasmányt a többi hallgatónak vagy oktatónak. Az ilyen ötletes és egyben szimbolikus dekorációk mind hozzájárulnak a többnyelvűség népszerűsítéséhez és fenntartásához.

Mivel a kutatás tíz hónapon keresztül tartott, így lehetőségünk nyílt több időszakos kiállítást is megfigyelni. Például az Irka gyermeklap harmincadik évfordulójára kiállítást szerveztek az eddig megjelent lapszámok illusztrációiból, és minden Irka-füzet hátoldalán szerepelt magyarul és ukránul az összes fontos információ az adott lapszámról. Egy másik kiállítás alkalmával a málenkij robot 80. évfordulójáról emlékeztek meg kétnyelvű (magyar–ukrán és magyar–angol) informatív poszterkiállítás keretében. Halloween alkalmából pedig annak lehettünk szemtanúi, hogy a Hallgatói Önkormányzat tagjai angol nyelvű dekorációkat hoztak létre az ünnepkör angolszász eredete miatt (4. ábra).

4. ábra. Időszakos többnyelvű kiállítások és dekorációk



6.2. Szimbolikus feliratok, jelek

A szimbolikus feliratok szobrokon és emléktáblákon jelentek meg. A Rákóczi-főiskolán található szobrok kivétel nélkül mind a magyar történelem nagyjait ábrázolták (pl. az intézmény névadóját, II. Rákóczi Ferencet), vagy a kárpátaljai magyarság jeles képviselőit (pl. Soós Kálmánt, az intézmény korábbi rektorát). Ezek a szobrok mind kizárólag magyar nyelvű feliratokat tartalmaztak.

Ha életrajzi adatok indokolták, akkor kétnyelvű feliratokkal látták el az emléktáblákat. Például Milton Friedman magyar–angol nyelvű emléktáblája esetében, aki beregszászi magyar szülők gyermeke volt, de miután szülei kivándoroltak Amerikába, ott lett belőle híres közgazdász. Egy másik példa Zupkó József, aki beregszászi magyarként lett Ukrajna érdemes kézilabdaedzője, így magyar–ukrán kétnyelvű emléktáblával tisztelgették előtte (5. ábra).

5. ábra. Szobor magyar felirattal és kétnyelvű emléktábla



Ami a nemzeti szimbólumokat illeti, Ukrajna zászlaja, címere és himnusza nagy hangsúlyt kapott az intézmény nyelvi tájképében. Egy esetben találkozhattunk olyannal, hogy az ukrán és a magyar címer és zászló egymás mellett lógott, a többi esetben a magyar nemzeti szimbólumok csak az emlékszobrok és -táblák melletti koszorúkon voltak megfigyelhetők piros–fehér–zöld szalagok formájában (6. ábra). Tehát bár az informális dekorációk és szimbolikus feliratok körében többnyire a magyar nyelv dominált, a nemzeti szimbólumok leginkább az intézménynek otthont adó ország szimbólumaira korlátozódtak.

6. ábra. Ukrán és magyar nemzeti szimbólumok



6.3. Hirdetmények

A hirdetményeket belső és külső típusokra osztottuk. A belső hirdetményeket maga az intézmény tette közzé, amelyeken valamilyen tudományos rendezvényt, konferenciát, ösztöndíjat népszerűsítettek. Ide tartoztak még a felvételi kampányhoz kapcsolódó plakátok is. A legtöbb hirdetmény kétnyelvű volt, amit két technikával értek el: vagy egy plakáton tüntették fel ugyanazt az információt magyar és ukrán nyelven, vagy ugyanabból a plakátból készítettek egy magyar és egy ukrán nyelvű változatot, egymás mellé vagy alá ragasztva őket. Az ilyen típusú kétnyelvű jelzéseket nevezi Reh (2004) duplikálónak.

7. ábra. Kétnyelvű belső hirdetőanyagok



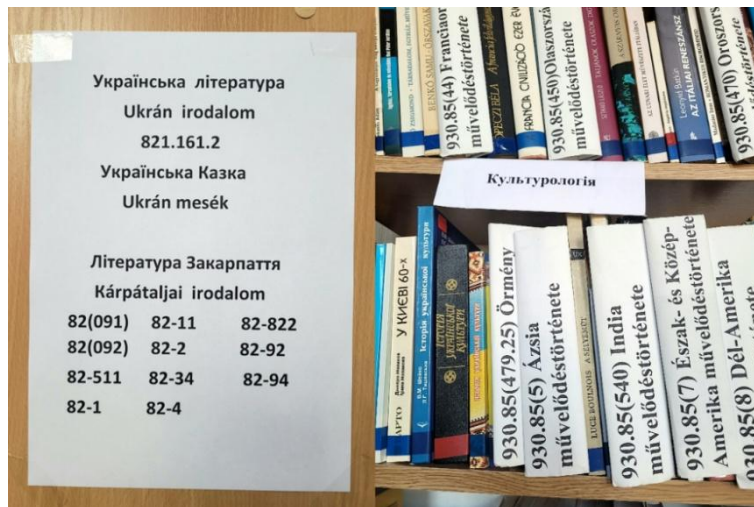
A főiskolai felvételit hirdető plakátokat csak magyar nyelven helyezték el, a weboldalon viszont ukránul is elérhetőek voltak. A főiskola folyosóin található hirdetőtáblákon a felvételi kampányt tekintve egyedül a Filológia Tanszék által készített magyar és ukrán nyelvű, QR-kóddal ellátott tájékoztató anyagok voltak kétnyelvűek (7. ábra).

A külső hirdetőanyagok közé főként álláshirdetések tartoztak, amelyek egy- és kétnyelvűsége annak függvényében változott, hogy az ideális munkavállalótól melyik nyelv(ek) ismeretét várták el. Ha a magyar és az ukrán nyelv ismerete is elvárás volt, akkor a hirdetés is két nyelven tartalmazott információkat. Kizárólag magyar nyelvűek voltak a kiadó lakások hirdetése és a kulturális szervezetek által szervezett koncertek vagy jótékonysági estek hirdetőanyagai.

6.4. Könyvtári jelzések

Általánosan elmondható, hogy a könyvespolcokon változatosan jelentek meg a magyar és az ukrán jelzések. A könyvek tudományterületeit vagy magyar–ukrán kétnyelvű feliratokkal tüntetik fel, vagy az adott polc könyveinek fő tudományterületét csak ukránul, míg a specifikusabb részterületeket csak magyarul (8. ábra). A könyvtárakban emellett számos magyar plakát emlékeztette a hallgatókat arra, hogy a főiskolán magyarországi tudományos adatbázishoz férhetnek hozzá ingyenesen.

8. ábra. Kétnyelvű könyvtári jelzések



A magyar nyelv dominanciája figyelhető meg a könyvtári olvasójegyen, amelyen bár a főiskola neve és az „olvasójegy” kifejezés szerepel a magyar mellett ukránul is, az összes többi információ mégis csak magyar nyelven olvasható, beleértve a könyvkölcsönzés szabályait is (9. ábra).

9. ábra. Könyvtári olvasójegy

Бібліотека Закарпатського угорського інституту ім. Ференца Ракоці II
II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola Könyvtára
90200 Beregszász, Kossuth tér 6.
@: kölcsönzo@kmtf.uz.ua

ЧИТАЦЬКИЙ КВИТОК
№ 543 számu olvasójegy
Váradi Krisztián
részére
Tanszék: Filológia Tanszék

ellenjegyzés

Fénykép helye
3*4 cm

Az olvasótermek látogatására is jogosít

Tudnivalók a kölcsönzéshez

1. Főállású oktatónak egyszerre 10 (másodállású 5 db) könyv kölcsönözhető.
2. A kölcsönzési idő: 2 nap az egy példányos, 2 hónap a többire.
3. A kölcsönzési idő túllépéséért késedelmi díjat köteles fizetni.
4. Amíg tartozását nem egyenlíti ki, addig nem kölcsönözhet.
5. Olvasójegy nélkül nem kölcsönözhet.
6. Az elveszett olvasójegy helyett újat pótdíj ellenében állítunk ki.
7. Az Ön olvasójegyével kikölcsönzött könyvek listája ellen kifogása nem lehet! Az olvasójegy száma az Ön azonosítója, vigyázzon, hogy visszaélés ne történjen!

A fentieket tudomásul veszem:

A könyvtár szépirodalmi kölcsönző részlege rendszeresen tart irodalmi kiállításokat híres magyar és ukrán költők és írók műveiből. Márciusban például egy külön polcon állították ki Tarasz Sevcenko legfontosabb műveit és a költő portréját születésnapja alkalmából.

6.5. Krízishelyzetekhez kapcsolódó kiírások

Ide tartoztak a tüzesetek, amelyekkel kapcsolatban számos egynyelvű ukrán biztonsági előírás volt a falakon. A menekülési tervek mind ukrán–magyar kétnyelvűek voltak, így minden oktató és hallgató könnyen megtalálhatja a legközelebbi kijáratot vészhelyzet esetén. Ebbe a kategóriába soroltuk a COVID-19 járvánnyal kapcsolatos útmutatásokat, amelyek vegyesen voltak egy- és kétnyelvűek is. Az intézmény személyzete által készített, kézfertőtlenítésre vagy maszkhasználatra figyelmeztető jelzések mind ukrán–magyar kétnyelvűek voltak, elsőként az ukrán nyelvű információt feltüntetve (10. ábra).

10. ábra. Ukrán–magyar kétnyelvű COVID-19 jelzések



Ebbe a kategóriába soroltuk az orosz–ukrán háború kapcsán a légvédelmi óvóhelyek felé vezető utat jelző nyilakat, valamint az óvóhelyek férőhelyeit jelző táblákat. Ezek mind ukrán–magyar kétnyelvűek voltak, ismét az ukrán szöveget feltüntetve elsőként. Az intézmény egyik volt hallgatója 2014-ben vesztette életét a Kelet-Ukrajnában zajló katonai hadműveletekben. Az ő arcképe egy ukrán zászló kíséretében azóta is ki van állítva a főiskola folyosóján (11. ábra).

11. ábra. Az orosz–ukrán háború hatása a főiskola nyelvi tájképére



6.6. Teremnevek

A főiskolán található osztálytermek, tanszékek, irodák, kiállítások és könyvtárak mind rendelkeznek kétnyelvű névtáblával, amelyen magyar–ukrán nyelven szerepel az adott helyiség megnevezése, az ott dolgozó személyek neve, illetve egyes esetekben, például könyvtáraknál a nyitvatartási idő is. Ezt szintén két időzónában tüntettek fel: a Magyarországon használt közép-európai idő és az Ukrajnában használt kelet-európai idő szerint (12. ábra).

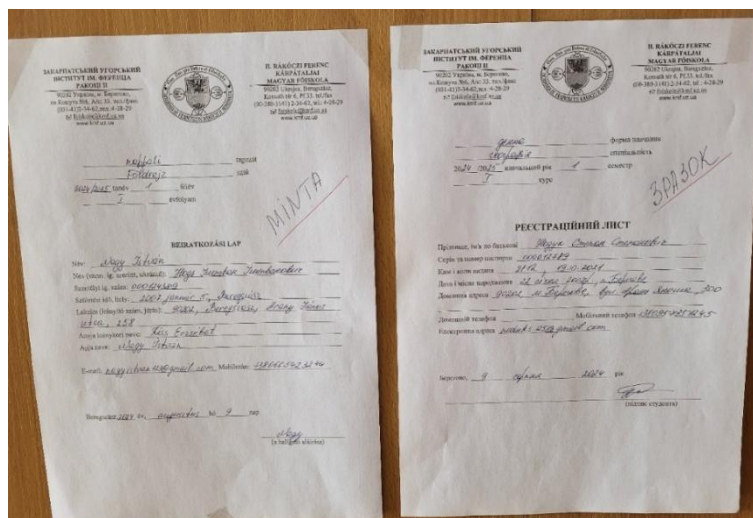
12. ábra. Kétnyelvű teremnévtáblák



6.7. Oktatási dokumentumok

Az oktatási dokumentumok már egyértelműen a fentről lefelé irányuló szabályozás kategóriájába esnek. A vizsgált intézményben gyakorlatilag minden oktatással kapcsolatos írásos dokumentum elérhető volt magyar és ukrán nyelven egyaránt, beleértve az órarendeket, a félévszerkezeteket, a szakdolgozati témalistákat, a választható tantárgyak listáját, a félévek elején kitöltendő beiratkozási lapokat, valamint a tanszékeken dolgozó oktatók fogadóóráit (13. ábra).

13. ábra. Kétnyelvű beiratkozási lapok



A kárpátaljai többnyelvű nyelvi tájkép egyik kitűnő példája az Angol Tanszék munkatársainak fogadóóráit bemutató táblázat, ahol három nyelv keveredését figyelhetjük meg: a dokumentum neve, a napok és az időpontok mind angol nyelvűek, míg a tanárok nevei ukránul és latin betűs átírással is fel vannak tüntetve, a *tanszékvezető* mint beosztás pedig magyar nyelven szerepel a lap alján (14. ábra).

14. ábra. Három nyelv keveredése

OFFICE HOURS PHILOLOGY DEPARTMENT ENGLISH SUBDIVISION							
		Monday	Tuesday	Wednesda y	Thursday	Friday	Online
1	Баняк Н. Ю. /Banyas N.	3-5pm	3-5pm				3-5pm
2	Баняк В. В. /Banyas V.	3-5pm		3-5pm			3-5pm
3	Барань А. Б. /Barany A.	3-5pm			3-5pm		3-5pm
4	Баті С. З. /Baty S.		3-5pm	3-5pm			3-5pm
5	Врабель Т. Т. /Vrabely T.	3-5pm		3-5pm			
6	Гнатик К. Б. /Hnatik K.			3-5pm	3-5pm		
7	Горенко О. П. /Horenko O.	3-5pm				3-5pm	3-5pm
8	Густі І. І. /Husti I.	3-5pm		3-5pm			
9	Деметська В. В. /Demetskaya V.						
10	Леврінц М. І. /Levrints M.		3-5pm		3-5pm		
11	Лехнер І. І. /Lchner I.	3-5pm			3-5pm		
12	Лівак К. М. /Lizak K.	3-5pm			3-5pm		
13	Надь-Колозшари Е. А. / Nagy-Kolozshvari		3-5pm	3-5pm			
14	Пліска В. Т. /Pliska V.			3-5pm	3-5pm		
14	Сіладі В. В. /Silagyi L.		3-5pm		3-5pm		
15	Фабіан М. Ю. / Fabian M.			3-5pm	3-5pm		
16	Ференц В. С. /Ferents V.	3-5pm		3-5pm			
17	Фодор К. Й. /Fodor K.		3-5pm		3-5pm		

Tanszékvezető: Beregsci A.Ф./Beregsczi A.

6.8. Intézményi feliratok

Minden intézményi névtábla ukrán–magyar kétnyelvű volt, mivel felépítésüket Ukrajna törvényei szabályozzák. Ezek előírják, hogy az intézmény nevének először ukrán nyelven kell szerepelnie, és csak alatta vagy tőle jobbra lehet elhelyezni angolul vagy az oktatás nyelvén (Karmacsi, 2018) (15. ábra).

15. ábra. Kétnyelvű intézményi névtáblák



6.9. Egyéb nyelvi jelek

Az egyéb nyelvi jelek közé tartoztak azok a nyelvitájkép-elemek, amelyeket nem tudtunk egyértelműen besorolni egyik specifikus csoportba sem. Itt egyaránt megtalálhatók informális és formális jelzések, ezért helyezkedik el az alulról felfelé és a fentről lefelé irányuló szabályozás skáláján (1. ábra) a középső helyen ez a kategória.

Az alulról felfelé irányuló informális jelzések döntő többsége magyar egynyelvű volt. Ezeket általában az intézmény munkatársai ragasztották ki, és olyan információkat közöltek, mint például a portás telefonszáma vagy a nyári nyitvatartási idő a könyvtárakban (16. ábra). Voltak olyan esetek is, amikor a kutatás első felében egy nyelvi jelet csak magyarul lehetett megfigyelni, majd ősszel már angolul is elérhető volt a kiírás, például a wifi-jelszó esetén.

16. ábra. Egynyelvű informális kiírások



A nem hivatalos jelzések közé tartoztak még a két főiskolai menzával kapcsolatos kiírások, például a napi menük. Ez mindkét adatolt esetben csak magyar nyelven volt elérhető, mivel a menzák nyelvhasználatát semmilyen

törvény nem szabályozza (17. ábra). Így ezek is alulról felfelé irányuló nyelvi jeleknek minősülnek.

17. ábra. Egynyelvű menük



A fentről lefelé irányuló formális nyelvitájkép-elemek közé főleg általános jellegű informatív jelzések tartoztak, mint például a kijáratokat, a lifteket, és az illemhelyeket jelző táblák (18. ábra). Általában ezek ukránul és/vagy angolul jelentek meg, bár a vonatkozó angol szavak a globalizáció hatására már mindenki számára ismertek. Például: *Exit*, *Toilets*, *Disabled*, *Men*, *Women*. Az esetek felében a magyar nyelv is feltűnt a táblákon az ukrán felirat mellett.

18. ábra. Angol–ukrán–magyar táblák



Ami a további hivatalos nyelvitájkép-elemeket illeti, különböző jogi dokumentumok is láthatók voltak a tanszékek melletti hirdetőtáblákon és a Tanulmányi Osztályon. Ezek mind csak ukrán nyelven voltak elérhetők. Ide tartoztak a szerződésminták, az intézményen belüli magatartási szabályok, illetve a kötelező katonai nyilvántartásba vételről szóló rendeletek.

6.10. Virtuális iskolai nyelvi tájkép

A virtuális nyelvitájkép-elemzés betekintést enged összesen hat, a főiskolához és annak részegységeihez kapcsolódó weboldal nyelvhasználatába, ezzel kitérve a Rákóczi-főiskola nyelvi tájképének megastruktúrájára. Az elemzés magába foglalta a főiskola weboldalát⁴, publikációtárát⁵, az intézmény Facebook-⁶, Instagram-⁷ és TikTok-oldalait⁸, valamint a Kiscsikó menza Facebook-oldalát⁹.

A főiskola háromnyelvű weboldallal rendelkezik, ahol minden új hír elolvasható magyar, ukrán és angol nyelven egyaránt (19. ábra). Ez a többnyelvűség jellemző az intézmény Facebook-oldalára is, ahol mindhárom nyelven megosztják az újdonságokat. A főiskolához tartozó publikációtár szintén háromnyelvű (magyar, ukrán, angol), a könyvtári katalógus pedig ezek mellett elérhető románul és németül is.

19. ábra. Háromnyelvű hírek a főiskola weboldalán



A Rákóczi-főiskolának van egy saját fejlesztésű tanulásirányító rendszere, a magyar és ukrán nyelven elérhető IRIS, amit mindenki gond nélkül tud használni, anyanyelvtől függetlenül. Az IRIS-en belül a tanárok és a diákok digitális formában hozzáférhetnek kétnyelvű órarendekhez, vizsgarendekhez, egyéni tantervekhez, elektronikus leckekönyvekhez. A hallgatók kétnyelvű digitális diákigazolványt kapnak, a tanárok pedig kétnyelvű oktatói profilokat.

Bár az intézményi Facebook-oldalon szintén minden hírt magyar, ukrán és angol nyelven osztanak meg, az Instagram és a TikTok sokkal informálisabb közösségimédia-felületeknek számítanak, ahol a képek és a videók mellett csak magyar nyelvű leírásokkal találkozhatunk. A Facebook tehát úgy működik, mint egyfajta hivatalosabb jellegű közösségimédia-csatorna az intézmény számára, ahol fontosabb a többnyelvű kommunikáció, míg a másik két felület informálisabb. A Kiscsikó menza Facebook-oldalán az informális bejegyzések szintén kizárólag magyar nyelvűek, a legfontosabb kiírást, a napi menüajánlatot

⁴ <https://kmf.uz.ua/hu>

⁵ <https://dSPACE.kmf.uz.ua/jspui>

⁶ <https://www.facebook.com/IIRFKMF>

⁷ <https://www.instagram.com/rakoczi.foiskola>

⁸ <https://www.tiktok.com/@iirfkmf>

⁹ <https://www.facebook.com/kiscsikomenza>

viszont mindig közzéteszik magyar és ukrán nyelven is, szemben a 6.9. pontban említett táblás menüajánlattal (17. ábra).

7. Következtetések

Kutatásunk célja a II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola nyelvi tájképének a bemutatása volt, különös figyelmet fordítva az alulról felfelé és a fentről lefelé irányuló szabályozás (Ben-Rafael et al., 2006) alatt álló nyelvitájkép-elemek közötti nyelvi különbségekre, valamint a magyar, ukrán és angol nyelvek szimbolikus jelentőségére.

Az összegyűjtött 546 nyelvi jel és az elemzett 6 weboldal alapján kijelenthetjük, hogy a vizsgált intézményben a magyar nyelv volt dominánsabb az informális, alulról felfelé irányuló jelzéseken, például a szobrokon található szimbolikus feliratok és az intézmény munkatársai által létrehozott informatív kiírások esetében. Amikor a hallgatóknak vagy a főiskolai dolgozóknak volt választási lehetőségük, inkább a magyar nyelvet részesítették előnyben (pl. magyar nyelvű tablók és menzai menüajánlatok). Ezzel szemben a formálisabb, felülről lefelé irányuló jelzések a legtöbb esetben ukrán–magyar kétnyelvűek voltak, elsőként ukrán nyelven feltüntetve az információkat (pl. a COVID-19 járvánnyal és az orosz–ukrán háborúval kapcsolatos kiírások, menekülési tervek, intézményi névtáblák). A legformálisabb elemek (pl. jogi dokumentumok) viszont csak ukrán nyelven voltak elérhetők. A nemzeti szimbólumok közül pedig egyértelműen az ukrán zászló, himnusz és címer volt megfigyelhető több alkalommal.

A legtöbb adatolt nyelvi jel mégis magyar–ukrán kétnyelvű volt. Az intézményen belüli hirdetményeket általában magyar és ukrán nyelven ragasztották ki egymás mellé. A falakon olyan művészi kompozíciók voltak láthatók, amelyek az elfogadást hangsúlyozták a különböző anyanyelvű és kultúrájú emberek felé. Az osztálytermek, irodák és tanszékek névtáblája szintén kétnyelvű volt. Sőt, minden fontosabb oktatási dokumentum elérhető volt magyar és ukrán nyelven, fizikális és online formában is egyaránt.

Az angol nyelv használata korlátozódott az Angol Tanszék rendezvényeire és néhány általános információt közlő jelzésre, mint a kijáratot, az illemhelyeket vagy a wifi-jelszót jelző kiírások. Emellett minden hírt közzétesznek a főiskola weboldalán angolul is, így szólítva meg egy nagyobb célközönséget.

Tódor (2014a, 2014b, 2015) és Biró (2016) is úgy találta romániai kisebbségi oktatási intézmények kontextusában, hogy a felülről lefelé irányuló jelzések az állami hatóságok előírásainak megfelelően készültek el, míg az alulról felfelé ható nyelvitájkép-szférában magasabb volt a magyar nyelv használati gyakorisága. Ha a törvényi szabályozások a felülről jövő jelek esetén nem is teszik lehetővé, az alulról felfelé ható nyelvitájkép-elemeknek tükrözniük kell a többnyelvű régiók nyelvi sokszínűségét (Chali & Parapatics, 2022). Ez a Rákóczi-főiskola esetén a jelen vizsgálat tanúsága szerint kiemelten jellemző. Fontos azonban megjegyezni,

hogy megállapításaink nem általánosíthatók más intézményekre is, hiszen kutatásunkban csak egy konkrét magán intézmény nyelvi tájképét szemléltettük esettanulmány jelleggel.

Köszönetnyilvánítás

A tanulmány a Kulturális és Innovációs Minisztérium 2024-2.1.1-EKÖP kódszámú Egyetemi Kutatói Ösztöndíj Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.

Irodalom

- Balog F. (2023). *Két magyar tannyelvű iskola nyelvi tájképe a Rimaszombati járásban. Diplomamunka.* Komárom: Selye János Egyetem.
- Barni, M. & Bagna, C. (2015). The critical turn in LL: New methodologies and new items in LL. *Linguistic Landscape*, 1(1–2), 6–18. <https://doi.org/10.1075/ll.1.1-2.01bar>
- Bartha Cs., Laihonon, Petteri & Szabó T. P. (2013). Nyelvi tájkép kisebbségben és többségben egy új kutatási területről. *Pro Minoritate*, 22(3), 13–28. Letöltés: <http://real.mtak.hu/7158>
- Ben-Rafael, E., Shohamy, E., Amara, M. H. & Trumper-Hecht, N. (2006). Linguistic landscape as symbolic construction of the public space: The case of Israel. *International Journal of Multilingualism*, 3(1), 7–30. <https://doi.org/10.1080/14790710608668383>
- Biró, E. (2016). Learning schoolscapes in a minority setting. *Acta Universitatis Sapientiae, Philologica*, 8(2), 109–121. <https://doi.org/10.1515/ausp-2016-0021>
- Biró, E. (2018). More than a Facebook share: Exploring virtual linguistic landscape. *Acta Universitatis Sapientiae, Philologica*, 10(2), 181–192. <https://doi.org/10.2478/ausp-2018-0022>
- Brown, K. D. (2005). Estonian schoolscapes and the marginalization of regional identity in education. *European Education*, 37(3), 78–89. <https://doi.org/10.1080/10564934.2005.11042390>
- Brown, K. D. (2012). The linguistic landscape of educational spaces: Language revitalization and schools in Southeastern Estonia. In Gorter, D., Marten, H. F. & Van Mensel, L. (eds.), *Minority Languages in the Linguistic Landscape* (281–298). London: Palgrave Macmillan. https://doi.org/10.1057/9780230360235_16
- Cenoz, J. & Gorter, D. (2006). Linguistic landscape and minority languages. *International Journal of Multilingualism*, 3(1), 67–80. <https://doi.org/10.1080/14790710608668386>
- Chali, K. K. & Parapatics, A. (2022). The Linguistic Landscape of Commercial Shop Signs of Nekemte town, Oromiya, Ethiopia. *Alkalmazott Nyelvtudomány*, 22(2), 89–104. <https://doi.org/10.18460/ANY.2022.2.006>
- Csernicskó I. (2015). Nyelvek vetélkedése a nyelvi tájképben: kárpátaljai példa. *Alkalmazott Nyelvtudomány*, 15(1–2), 71–84. Letöltés: <http://real.mtak.hu/31648>
- Csernicskó I. (2016). A változás megragadása a nyelvi tájképben: Kárpátalja példája. *Magyar Nyelv*, 112(1), 50–62. <https://doi.org/10.18349/MagyarNyelv.2016.1.50>
- Csernicskó I. (2022). Nyelvi (emberi) jogok, jogsértések és biztonságpolitika összefüggései Oroszország Ukrajna elleni háborúja tükrében. *Fórum Társadalomtudományi Szemle*, 24(2), 3–30. Letöltés: <https://real.mtak.hu/144859>
- Csernicskó I. & Gazdag V. (2023). Kárpátaljai magyarok nyelvi tapasztalatai Magyarországon. *Magyar Nyelv*, 119(2), 215–225. <https://doi.org/10.18349/MagyarNyelv.2023.2.215>
- Csernicskó I. & Tóth M. (2021). *A kárpátaljai magyar közösség nemzetiségi és nyelvi jogai. Kárpát-medencei magyar kisebbségjogi kalauz.* Budapest: Bethlen Gábor Alapkezelő Zrt. Letöltés: <http://real.mtak.hu/137964>
- Csernicskó, I., Hires-László, K., Karmacs, Z., Márku, A., Molnár D., I., Máté, R. & Tóth-Orosz, E. (2023). *Hungarians and the Hungarian Language in Transcarpathia.* Törökbálint: Termini Egyesület. Letöltés: <https://real.mtak.hu/162720>

- Fedinec Cs. (2024). A nemzeti kisebbségekkel kapcsolatos ukrainai jogalkotás rövid foglalata (1989–2023). *Fórum Társadalomtudományi Szemle*, 26(4), 65–80. <https://doi.org/10.61795/fssr.v26y2024i4.04>
- Gorter, D. & Cenoz, J. (2015). Linguistic landscapes inside multilingual schools. In Spolsky, B., Tannenbaum, M. & Inbar, O. (eds.), *Challenges for Language Education and Policy: Making Space for People* (151–169). New York: Routledge.
- Hires-László, K. (2019). Linguistic landscapes in a Western Ukrainian town. *Corvinus Journal of Sociology and Social Policy*, 10(1), 87–111. <https://doi.org/10.14267/CJSSP.2019.1.5>
- Hires-László, K., Máté, R. & Tóth-Orosz, E. (2023). Multilingual COVID-19 linguistic landscape in Berehove. *Philological Review*, 21(1), 5–17. <https://doi.org/10.31499/2415-8828.1.2023.281351>
- Istók B., Lőrincz G. & Török T. (2024). Módszertani megfontolások az iskolai és óvodai nyelvi tájkép kutatásához. In Polgár A., Keserű J., Lőrincz G. & H. Nagy P. (szerk.), *Didicisse fideliter artes. Köszöntő kötet Szabó András tiszteletére* (235–262). Komárom: Selye János Egyetem.
- Johnson, N. B. (1980). The material culture of public school classroom: The symbolic integration of local schools and national culture. *Anthropology & Education Quarterly*, 11(3), 173–190. <https://doi.org/10.1525/aeq.1980.11.3.05>
- Karmacsi Z. (2018). Kárpátaljai magyar tannyelvű iskolák nyelvi tájképe. In Karmacsi Zoltán & Máté Réka (szerk.), *Nyelvek és nyelvváltozatok térben és időben. Tanulmányok a Hodinka Antal Nyelvészeti Kutatóközpont kutatásaiból IV.* (174–182). Ungvár: „RIK-U”.
- Karmacsi Z. & Mészár K. (2024). Kárpátalja egyházi fenntartású iskoláinak virtuális nyelvi tájképe. *Acta Academiae Beregsasiensis, Philologica*, 3(2), 90–108. <https://doi.org/10.58423/2786-6726/2024-2-90-108>
- Laihonen, P. & Tódor, E.-M. (2017). The changing schoolscape in a Szekler village in Romania: signs of diversity in rehungarization. *International Journal of Bilingual Education and Bilingualism*, 20(3), 362–379. <https://doi.org/10.1080/13670050.2015.1051943>
- Landry, R. & Bourhis, R. Y. (1997). Linguistic landscape and ethnolinguistic vitality: An empirical study. *Journal of Language and Social Psychology*, 16(1), 23–49. <https://doi.org/10.1177/0261927X970161002>
- Lőrincz G. (2021). Identitásjelölő elemek egy délnyugat-szlovákiai kisváros oktatási intézményeinek nyelvi tájképében. *Acta Universitatis de Carolo Eszterházy Nominatae. Sectio Linguistica Hungarica*, 47, 61–71. <https://doi.org/10.46437/ActaUnivEszterhazyLinguistica.2021.61>
- Lőrincz G., Istók B. & N. Varagya Sz. (2021). A weboldalak szerepe a virtuális iskolai nyelvi tájképben. *Hungarológiai Közlemények*, 22(1), 84–95. <https://doi.org/10.19090/hk.2021.1.84-95>
- Molnár J. & Molnár D. I. (2005). *Kárpátalja népessége és magyarsága a népszámlálási és népmozgalmi adatok tükrében*. Beregszász: A KMPSZ Tankönyv- és Taneszköz Tanácsa.
- Ohár B. (2022). *Iskolai nyelvi tájkép Kárpátalján: A Beregszászi Bethlen Gábor Magyar Líceum, a Nagyberegi Református Líceum és a Nagyberegi Dobrai Péter Középiskola nyelvi tájképének összehasonlító elemzése. Diplomadolgozat*. Beregszász: II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola. Letöltés: <https://dspace.kmf.uz.ua/jspui/handle/123456789/1863>
- Reh, M. (2004). Multilingual writing: A reader-oriented typology – with examples from Lira Municipality (Uganda). *International Journal of the Sociology of Language*, 170, 1–41. <https://doi.org/10.1515/ijsl.2004.2004.170.1>
- Shohamy, E. & Ben-Rafael, E. (2015). Introduction. Linguistic Landscape: A new journal. *Linguistic Landscape*, 1(1–2), 1–5. <https://doi.org/10.1075/ll.1.1.2.001int>
- Szabó, T. P. (2015). The management of diversity in schools: An analysis of Hungarian practices. *Apples – Journal of Applied Language Studies*, 9(1), 23–51. <https://doi.org/10.17011/apples/2015090102>
- Tódor E.-M. (2014a). Hogyan hívogat az iskola? Vizuális nyelvhasználat az iskolában. *Magiszter*, 12(4), 40–51.
- Tódor, E.-M. (2014b). The hidden curriculum of schools: overview of the bilingual school context. *Journal of Romanian Literary Studies*, 4, 529–538.
- Tódor E.-M. (2015). Kie az iskola? Nyelvi tájkép és nyelvhasználati szokások a székelyföldi magyar tannyelvű iskolákban. In Benő Attila, Fazakas Emese & Zsemlyei Borbála (szerk.), *Többnyelvűség és kommunikáció Kelet-Közép-Európában* (256–265). Kolozsvár: Erdélyi Múzeum Egyesület.

- Tótor E.-M. (2019). *Hétköznapí kétnyelvűség. Nyelvhaználát, iskolai nyelvi tájkép és nyelvi én a romániai magyar iskolákban*. Budapest: Ráció Kiadó és Szépirodalmi Figyelő Alapítvány.
- Tóth M. (2024). Ukrajna európai uniós integrációja és a nemzeti kisebbségek jogainak problémája. *Kisebbségvédelem*, 9, 41–98.
- Tóth-Orosz E. (2020). A beregszászi járási önkormányzatok virtuális nyelvi tájképe. *LIMES: A II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola tudományos évkönyve*, 7(2), 199–208. <https://dspace.kmf.uz.ua/jspui/handle/123456789/2006>
- Váradí K. (2023). Kétnyelvűséghez fűződő attitűdök vizsgálata többnyelvű környezetben élő főiskolai filológus hallgatók körében. *Argumentum*, 19, 88–123. <https://doi.org/10.34103/ARGUMENTUM/2023/6>
- Váradí K. (2025). Az iskolai nyelvitájkép-kutatások elméleti és módszertani háttere. *Acta Academiae Beregsasiensis, Philologica*, 4(2) (megjelenés alatt).
- Váradí, K., Hires-László, K. & Máté, R. (2024). Schoolscaping in a multilingual area: An example of a Hungarian college in Western Ukraine. *Sustainable Multilingualism*, 25, 165–195. <https://doi.org/10.2478/sm-2024-0016>
- Veres Cs. (2024). *Beregszász és a Beregszászi kistérség iskoláinak „virtuális” nyelvi tájképe. Diplomadolgozat*. Beregszász: II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola. Letöltés: <https://dspace.kmf.uz.ua/jspui/handle/123456789/4180>
- Wedin, A. (2021). Schoolscaping in the third space: The case of the Language Introduction Programme in Sweden. *Apples – Journal of Applied Language Studies*, 15(2), 151–168. <https://doi.org/10.47862/apples.110860>

Juhász Kornélia – Deme Andrea: A helyesírásban jelölt és nem jelölt *j* akusztikai megvalósulása gyors beszédtempóban

Alkalmazott Nyelvtudomány, XXV. évfolyam, 2025/1. szám, 24–43.

doi:<http://dx.doi.org/10.18460/ANY.2025.1.002>

A helyesírásban jelölt és nem jelölt *j* akusztikai megvalósulása gyors beszédtempóban

The study examines how the speed of speech tempo affects the production of Hungarian *j*-realizations in words where the *j* element is either represented in Hungarian orthography or not. Specifically, the experiment examines the acoustic discrimination of word pairs involving inflected infinitives and nouns. For example, *zoknija* [zoknijɒ] ('his socks') contains a possessive suffix, where the *j* is both orthographically represented and morphologically confirmed as lexical. In contrast, *szoknia* [soknijɒ] ('to get used to') is an inflected infinitive in the 3rd person singular, where the absence of *j* may be attributed merely to orthographic convention rather than morphological evidence. Although the acoustic discrimination of the *ija* and *ia* sequences of these words is not phonologically motivated, since *j* is not contrastive in this context in Hungarian, previous studies showed that speakers produce a spectral contrast between them at a comfortable speech rate, possibly as a result of the effect of orthography. Specifically, the *ija* sequence is produced with a more constricted vocal tract during the transitional phase between the [i] and [ɒ] vocalic targets, resulting in a lower first formant frequency (F₁), compared to the *ia* sequence. However, the duration of the vocalic sequences was not affected by the orthographic representation. It is well established that fast speech induces segmental reduction due to temporal constraints. Based on this, in the present study, we hypothesized that while orthography-driven distinctions influence speech at a comfortable rate, spectral differences between *ija* and *ia* disappear in fast speech.

To test this, we recorded 15 female speakers producing four phonetically matched real words in carrier sentences (*Andija* vs. *adnia*; *zoknija* vs. *szoknia*; *haknija* vs. *raknia*; *pónija* vs. *rónia*). These words were first read at a comfortable speech rate (normal tempo) in carrier sentences (*Legyen <target word>! 'let it be <target word>'*), yielding the results described above. Subsequently, speakers were asked to repeat these sentences multiple times, gradually increasing their speech tempo until articulation failed to proceed. This method facilitated both an increase in articulation rate and the automatization of articulatory gestures. We segmented the vocalic sequences of words (marked in bold) manually and measured both the duration of the target vocalic sequence and F₁ at 5 ms intervals throughout its total duration. Duration data were compared by GLMM, while F₁ was analysed using GAMMs.

The results confirmed that the *ija* and *ia* sequences did not differ in duration, even at fast speech rates, consistent with the absence of temporal contrast observed at normal speech rates in previous studies. Regarding spectral characteristics, it was hypothesized that faster speech would neutralize the differences between the sequences along the dimension of the closeness of the vocal tract. However, the differences observed in normal speech rate persisted at fast rate: the *ija* sequence exhibited a more constricted vocal tract during the transitional phase, as reflected in lower F₁ values, maintaining its

distinction from *ia*. Interestingly, fast speech reduced the [i] vowel in both *ija* and *ia* to an approximant-like realization, while the [ɒ] vowel remained relatively resistant to reduction effects exerted by increased speech rate. These findings contribute to our deeper understanding of how orthographic representation influences speech production and how Hungarian *j* realizations are produced.

Keywords: Hungarian *j* realizations, fast speech tempo, speech production, coarticulation, effect of orthography

1. Bevezetés

A jelen tanulmányban bemutatott kísérlet célja annak vizsgálata, hogy hogyan hat a beszédsebesség felgyorsítása a helyesírás által megjelenített és nem megjelenített *j*¹-megvalósulások ejtésére valódi magyar szavakban, emellett a felgyorsított tempójú megvalósulásokat a kényelmes (a későbbiekben: „normál”) tempójú realizációkkal is összevetjük. A vizsgálat középpontjában két olyan *j*-realizáció megvalósulása áll, amelyek a magyar nyelvben fonológiailag nem állnak kontrasztban, tehát nem rendelődik hozzájuk jelentésmegkülönböztető szerep. Éppen ezért ezeknek az elemeknek az akusztikai elkülönítése – a fonológiai rendszer szempontjából – nem motivált. Ennek ellenére korábbi kísérletekben azt tapasztalták, hogy a beszélők normál beszédtempóban spektrális tulajdonságaikban mégis elkülönítik a helyesírásban jelölt, illetve nem jelölt *j*-t tartalmazó hangsorokat, ami elsősorban valószínűsíthetően az ortográfia, azaz az íráskép hatásának tulajdonítható. Habár a jelen kísérlet az ortográfia facilitálta ejtés jellemzőit tárgyalja, a kísérlet fő kérdése mögött, és e kísérlet előzményeként egy terjedelmes fonológiai szakirodalom is áll (pl. Siptár & Törkenczy, 2000/2007; Siptár, 2002, 2005, 2011, 2013). A tanulmány legelején ezekre is támaszkodva mutatjuk be röviden az ortográfiai megjelenítés és a két *j*-realizáció megvalósulásának – feltételezett – eltéréseit.

A magyar ortográfia (kvázi-)fonematikusnak tekinthető: többé-kevésbé transzparens, azaz a legtöbb esetben közvetlen kapcsolat figyelhető meg a mögöttes fonéma, illetve az azt jelölő graféma között. A jelen tanulmány kérdésfelvetését illetően a *j* szegmentum helyesírási jelölése releváns. Amikor a helyesírásban *j*-t jelölünk (pl. *jel* vagy *nyáj*, akár *ly* grafémakombinációval is, pl. *ilyen*), akkor erről a szegmentumról azt feltételezzük, hogy jelen van a fonemikus reprezentációban. Amikor a *j* hiátustöltőként az ejtésben jelentkezik, de nem jelöli a helyesírás (pl. *dió*, *fiú*, *múmia*), akkor kérdéses, hogy része-e a fonemikus reprezentációnak. A magyar helyesírást e tekintetben ismerő beszélők számára feltehetően nem része (lásd alább részletesebben), azonban az írni-olvasni (még) nem tudó egyének számára a *dió j*-je éppen olyan státuszú szegmentum, mint *hájas* középső eleme (ezt olvasni tanuló első osztályos gyermekek rendre megerősítik, és a tanulási folyamat része annak az elvnek az elsajátítása, hogy a hallott és ejtett *j*-t a korábbi esetben ne tekintsék az írott forma részének). Ez

¹A tanulmányban *j*-vel (dőlt betűvel) utalunk a helyesírási konvencióra, illetve a kérdéses palatális elemekre általában is, státuszuktól függetlenül.

ugyanakkor természetesen azt is jelenti, hogy a fonemikus reprezentáció egyénfüggő, és függ a(z írott) nyelvi tapasztalattól, de erre vonatkozóan nincs biztos fogódzónk. Vagyis a hiátustöltő *j* meg nem jelenése az írásképből nem utal arra sem, hogy ez a *j* része a mentális fonémareprezentációnak, de arra sem, hogy nem az.

Siptár és Törkenczy (2000/2007) a magyar nyelvben két különböző *j*-realizációt feltételez, amelyek közötti eltérések úgy írhatók le, hogy a hiátustöltőként megvalósuló [j] fonetikai megvalósulása „gyengébb és átmenetibb jellegű”, mint a mögöttes /j/ realizációja (Siptár & Törkenczy, 2000/2007: 91). Ezzel párhuzamosan Siptár (2013), illetve Siptár és Törkenczy (2000/2007) a két *j*-realizáció absztrakt tulajdonságait is eltérőként mutatja be: míg a fonemikus /j/ egy mássalhangzós természetű likvida [+msh, +szon], addig a hiátustöltő [j] magánhangzós természetű siklóhang [-msh, +szon].

Ebben a kísérletben toldalékolt főnevek (pl. *zoknija*), illetve ragozott főnévi igenevek (pl. *szoknia*) félkövérrel jelölt magánhangzószerű elemeket tartalmazó (azaz vokalikus) részének vagy szakaszának megvalósulását elemezzük. Ezen példák esetében kizárólag az ortográfiai alakra hagyatkozhatunk, mert míg a *zoknija* esetében a *-ja* toldalékában mind morfológiailag, mind ortográfiailag jelen van a *j*, ezzel szemben a főnévi igenevek esetében a *j* graféma hiánya nem ad támpontot arra vonatkozóan, hogy része-e ez a szegmentum a mentális fonémareprezentációnak. A *szoknia* szó példájánál maradván a főnévi igenév toldalékának két allomorfja van: a *-ni* és az *-n*, viszont az egyes szám harmadik személyű személyragnak is ehhez hasonlóan kettő van: az *-a* és a *-ja*. Tehát ebből következően nem dönthető el, hogy az absztrakt fonémaszekvenciában a *ni* + *a* vagy a *ni* + *ja* allomorf-kombináció jelenik-e meg, hiszen a felszíni megvalósulásban mindkét esetben a [soknijɒ] megvalósulást kapjuk (Siptár, 2011: 155). A mögöttes formában az allomorfok további két kombinációjának megjelenése pedig kevésbé valószínű, hiszen a *-na* már „foglalt”, és ez az alak feltételes módú *szokna* [soknɒ] alakkal lenne homoním, míg a *-nja* fonotaktikailag és gyakoriság szempontjából tekinthető valószínűtlen választásnak (Siptár, 2011: 157, Rebrus & Kálmán, 2009). Összefoglalva: abból kiindulva, hogy a morfofonológiai elemzések a főnévi igenevek (pl. *szoknia*) esetében nem erősítik meg a *j* mögöttes jelenlétét/hiányát, ezeknek (és a hasonló) szópároknak az esetében kizárólag az lehet megalapozott kérdés, hogy találunk-e különbséget az ejtésben pusztán azért, mert a helyesírás eltérést sejtet a hangsorok között. A kísérlet anyagának összeállításakor azért döntöttünk ezeknek a szópároknak a vizsgálata mellett, mert az azokban található *ija* és *ia* szekvenciák a fonetikai vizsgálat szempontjából itt jelennek meg a lehető leginkább kontrollálható kontextusban, értve ezt mind a szón belüli pozíciójukra, a hangsúly-pozícióra és a beszédhang-környezetre is (vö. még Deme & Juhász, 2024).

Feltételezhetjük, hogy az ortográfia ismerete alapvetően befolyásolja és formálja az egyén elméjében lévő fonológiai rendszert, az egyén nyelvhasználatát,

észlelését és az ejtés jellemzőit is. Például olyan idegen eredetű magyar szavakban, amelyeknek az írásképe hosszú mássalhangzót tükröztet, előfordul annak hosszú ejtése, még akkor is, ha egyébként a rövid mássalhangzós ejtés az elterjedt, kodifikált (pl. *apparátus* [ɒppɒraːtuʃ] ~ [ɒpːɒraːtuʃ] vagy *akkord* [ɒkord] ~ [ɒkːord]), miközben olyan példákat is találunk, ahol egy rövid mássalhangzóval írt szó hosszú mássalhangzós ejtismódja az elterjedt, és „gondozott” ejtésben a rövid mássalhangzós ejtismód jelenik meg (pl. *csemláló* [ʃɛmbɒloː] ~ [ʃɛmbɒlːoː]) (Huszthy, 2023). Arra is ismerünk példákat, hogy az ortográfiának a nyelvtanulásban közvetlen hatása van mind a célnyelvi észlelésre, mind az ejtésre: például az angolt idegen nyelvként tanulók esetében gyakori a *debt* szóban a *b* mássalhangzó ejtése (vö. [debt]), szemben az anyanyelvű beszélők ejtésével, ahol az említett zárhang nem eleme az akusztikai/artikulációs kimenetnek (Bassetti, 2008). Hasonlóképpen például a kínai karakterek pinyin transzkripciója is – az ortográfiai hatás következtében – félrevezetheti a nyelvtanulókat. Szintén Bassetti (2008) eredményei szerint a kínaiul tanulók helyesen sajátították el a *you* [jou] rímet abban az esetben, amikor az ortográfiai leirat az összes akusztikai szerkezetben megjelenő szegmentumot megjelenítette (tehát az *ou* diftongust). Ugyanezen nyelvtanulók ugyanakkor nem valósították meg az [o] hangelemet a [jou] rímben akkor, ha az *o* graféma nem szerepelt az írásképből (pl. *liu* (pinyin) → *[liu]-ként ejtve, [liou] helyett). Más szóval a nyelvtanulók azokban esetekben nem ejtették az adott hangzót (illetve annak egy elemét), amikor az az ortográfiai megjelenítésben sem szerepelt.

A fentiekre alapozva a jelen tanulmányban abból indulhatunk ki, hogy az íráskép eltérése együtt járhat az ejtésben és az akusztikai szerkezetben jelentkező különbségekkel. Ez például azt jelentheti, hogy a *zoknija* és a *szoknia* szó helyesírásban megjelenő képe közötti eltérés az utolsó két szótag magánhangzószerű elemeket tartalmazó részében (*ija* vs. *ia*) eltérő ejtést és akusztikai szerkezetet hívhat elő, függetlenül attól, hogy a fonológiai reprezentációban mely elemek meglétét feltételezzük.

1.1 A kísérlet előzményei, valamint a vizsgált beszédhangsorok akusztikai szerkezete és tulajdonságai

A *j*-realizációk, mind fonológiai, mind fonetikai szempontú elemzése terjedelmes szakirodalommal rendelkezik a magyar nyelvre nézve, és ezek a jelen kísérlet előzményeit is képezik (pl. Menyhárt, 2006; Olaszy, 2010; Gósy, 2014a, 2014b, 2018). A *j*-változatok kutatásának korábbi irodalmát máshol tekintjük át részletesen (Juhász & Deme, 2022; Deme & Juhász, 2024). A jelen tanulmányban kizárólag két, a jelen vizsgálat szemszögéből leginkább közvetlennek ítélt kísérlet eredményeit mutatjuk be röviden. Ezekben a kísérletekben, a jelen vizsgálat módszereihez hasonlóan nem szegmentálták a szerzők a kérdéses *j* elemeket (szemben a korábbiakkal), és a teljes vokalikus szakasz időtartamát, valamint a toldalékcso zártságának akusztikai vetületét elemezték. Ráadásul

utóbbinak az elemzéséhez dinamikus módszert választottak a szerzők, ami szintén egyezik a jelen kísérlet megközelítésével.

Juhász és Deme (2024) tanulmányukban olyan álszavak ejtését vizsgálták, amelyekben *ia* és *ija* vokalikus szekvenciák dentális /s/, illetve posztalveoláris /ʃ/ szibilánskontextusban jelentek meg (vö. *sziá*, *szijá*, *siá*, *sijá*). Az *ija* és *ia* vokalikus szekvenciák ejtését álszavakban vizsgálva az ejtéséből közvetlenül következtetni lehet az ortográfiai hatásra, hiszen ebben az esetben a beszélők pusztán a megjelenített grafémasorokra támaszkodhattak a beszéd tervezésekor. E kísérlet eredményei megerősítették, hogy pusztán az ortográfiai megjelenítésre hagyatkozva a beszélők elkülönítik az *ija* és *ia* szekvenciákat kétféleképpen is. Egyfelől eltérés mutatkozott ezeknek az időtartamában, mégpedig úgy, hogy az *ija* vokalikus szakaszt hosszabb időtartammal ejtették az *ia*-hoz viszonyítva, ami közvetlenül – egyebek mellett – arra (is) utalhat, hogy az *ija* esetében eggyel több elem (beszédhang) jelentkezett a vokalikus szakaszban a megvalósításkor. Másfelől a beszélők spektrálisan is elkülönítették az álszavak *ija* és *ia* szekvenciáit: az *ija* vokalikus szakasz *j* realizációjának akusztikai vetületéből arra lehetett következtetni, hogy a toldalékcso az ejtés közben zártabb volt, vagyis a kísérleti személyek az *ia*-hoz képest mássalhangzószabban ejtették a vokalikus szakasz kérdéses szakaszát.

A fentebb említett vizsgálatban tehát álszavakat elemeztek a szerzők, amelyeknek az ejtésében csak a helyesírási kép lehetett a meghatározó, és amiknek az ejtésük fokozottan jelentkezhett a hiperartikulációra, azaz a beszédhangok minél pontosabb ejtésére irányuló törekvés (Maxwell et al., 2015) is a beszélőkben. A második itt idézendő elemzés ezekre az eredményekre épült, és ebben már valódi, jelentéssel rendelkező, és a jelen kísérletben is elemzett szavakban megvalósuló vokalikus szekvenciákat vizsgálták a szerzők szintén a jelen kísérletben részt vevő beszélők ejtésében: toldalékolt főneveket (pl. *zoknija*), illetve ragozott főnévi igeneveket (pl. *szoknia*) (Deme & Juhász, 2024). Deme és Juhász (2024) hipotézise szerint az írásképpen megjelenített, illetve nem megjelenített intervokális *j*-ket – az ortográfiai hatásnak köszönhetően – a magyar beszélők valódi szavak ejtésük is elkülönítik egymástól. A megelőző álszavas vizsgálattal egyezően ebben a kísérletben is a vokalikus szakaszok időtartamát és a toldalékcso zártságával összefüggésben álló akusztikai mérőszámot vizsgálták és vetették össze a szerzők az *ija* és az *ia* esetében. Ennek a kísérletnek az eredményeit a következőkben úgy mutatjuk be, hogy ezzel egyidőben kitérünk a vizsgált hangsorokban megvalósuló szegmentumok akusztikai tulajdonságainak ismertetésére is.

A fentebb grafémával *ia* és *ija*-ként megjelenített hangsorokban megvalósuló beszédhangokat általánosságban mind vokalikus, vagyis magánhangzószzerű szegmentumoknak tartjuk: míg az [i] és [ɒ] szegmentum magánhangzó, addig a (lexikálisan is jelentkező, prototipikus) [j] mássalhangzó, és azon belül is közelítőhang, idegen szóval approximáns. A vokalikus szegmentumok

képzésekor a levegő akadálymentesen áramlik át a toldalékcsovön. A toldalékcsov üregeinek méretét és formáját modulálva különböző akusztikai minőségű beszédhangok hozhatók létre abból következően, hogy a toldalékcsov alaki és méretbeli tulajdonságainak függvényében megszűri a gégében keletkezett hangot, vagyis a zöngét, azaz felerősít és gyengít bizonyos frekvenciájú összetevőket abban (Fant, 1960). A toldalékcsov gerjesztésekor felerősödő frekvenciákat (illetve a toldalékcsovra jellemző sajátrezonanciákat) nevezzük formánsoknak (Fant, 1960; Deme, 2016: 26; Juhász & Deme, 2022; Juhász, 2024a), és a vokalikus szegmentumokat egytől egyig formánsos akusztikai szerkezet jellemzi (Ladefoged, 1975).

A jelen kísérlet szempontjából a toldalékcsovben létrejövő szűkület mértéke az, ami elsődlegesen vizsgálandó akusztikai jegynek bizonyul. A fonetikai és fonológiai szakirodalom szerint a toldalékcsovben létrejövő szűkület minősége statikus szempontból eltér a [j] approximáns, illetve a magánhangzó-kapcsolatban megjelenő [j]-szerű hiátustöltő beszédhang ejtése, valamint az [i] ejtése között. Ez az eltérés „mássalhangzóság” mértékében, vagyis a toldalékcsov zártságának, a nyelvemelkedés fokának vagy a toldalékcsovben létrejövő szűkület keresztmetszete méretének a szempontjából jelentkezik. Röviden: míg az approximáns [j] a várakozások szerint mássalhangzósabb, ezért magasabb nyelvállás és kisebb szűkületi keresztmetszet jellemzi, a hiátustöltőként megjelenő [j] egy magánhangzósabb siklóhang, tehát alacsonyabb nyelvállással, nyíltabban, nagyobb szűkületi keresztmetszettel jön létre, amihez képest az [i] még inkább magánhangzós, tehát még alacsonyabb nyelvállással és még nyitottabban képződik (Cruttenden, 2008; Heselwood, 2006; Hunt, 2009; Jagers, 2018; Davidson & Erker, 2014; van Heuven & Hoos, 1991; részletesebben lásd még Juhász & Deme, 2022; Juhász & Deme, 2024; Deme & Juhász, 2024).

Mivel a várakozások szerint az /io/ magánhangzó-kapcsolatban a /j/ approximáns nem jelenik meg, pontosabban nem az approximáns jelenik meg, hanem egy siklóhang, ezért az sem várható, hogy a felső nyelvállású, zárt ejtésű [i] magánhangzónál zártabb toldalékcsov-keresztmetszet (vagyis magasabb nyelvállás) jelentkezne a két magánhangzós elem között (Heselwood, 2006).

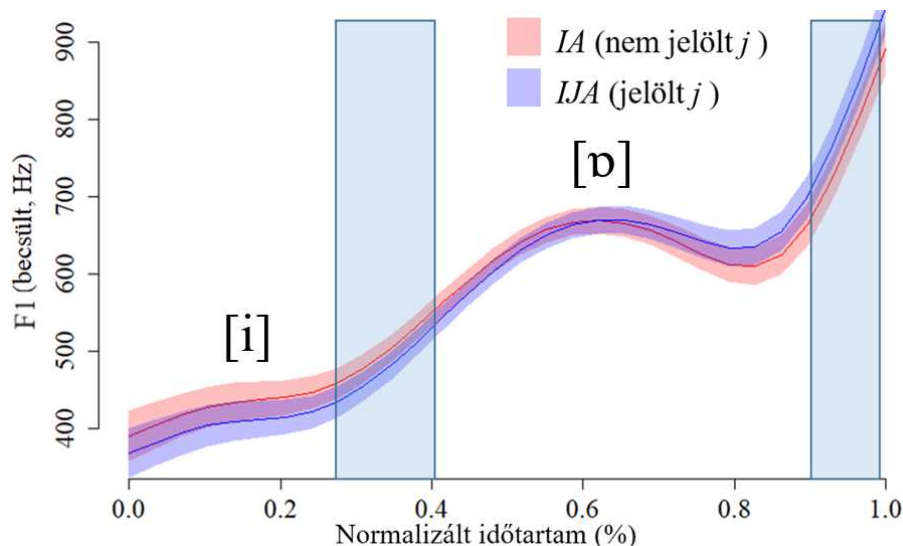
A nyelvemelkedés fokát, illetve a toldalékcsov zártságát akusztikailag az első formánsfrekvencia (F_1) értékével szokás számszerűsíteni, ami a nyelvemelkedés fokával fordított, a szűkület keresztmetszetével pedig egyenes arányban áll (Stevens, 2000). Ennek megfelelően a nyíltabb, magánhangzószerűbb hangokban (pl. az [v]-ban vagy [a]-ban) magasabb, a zártabb (pl. [i] és [j]) beszédhangokban pedig alacsonyabb F_1 -értékekre számíthatunk. Ugyanebből az összefüggésből kiindulva az [i] magánhangzó, illetve a [j] approximáns eltéréseit tekintve – még ha ezek az eltérések kizárólag tendencia szintjén jelennek is meg, amint azt Jagers (2018) eredményeiben látjuk – azt mondhatjuk, hogy a [j] approximánst az [i] magánhangzóhoz képest alacsonyabb F_1 -érték jellemzi a magasabb nyelvállásából fakadóan.

Fontos kiemelnünk, hogy a statikus nézőpontú, az ejtés egyetlen időpillanatára irányuló elemzés a [j] approximáns képzésének is csak egyetlen, pillanatnyi aspektusát jeleníti meg, miközben ismert, hogy a [j] approximánsnak az [i] magánhangzónál jelentősen dinamikusabb a megvalósulása. Ezt a dinamikus megvalósulást pedig érdemes a koartikuláció szempontjából bemutatni, így mindenekelőtt e fogalom tisztázására térünk ki. Ezzel párhuzamosan bemutatjuk az idézett kísérlet előzményeit (amik a jelen tanulmány alapjául szolgáltak) az *ija* és *ia* kényelmes beszédsebességű ejtésére vonatkozóan.

Az emberi beszéd fiziológiai alapja az, hogy a beszédszervek folyamatosan mozgásban vannak, vagyis az időben előre haladva újabb és újabb artikulációs-akusztikai célokat közelítenek meg, érnek el, majd hagynak maguk mögött. A beszéd ilyen folyamatosan változó természetét nevezzük koartikulációnak, ez az elképzelés pedig magában foglalja azt a feltételezést, hogy (i) az egyes beszédhangokhoz önálló artikulációs-akusztikai tartoznak, valamint azt is, hogy (ii) ezek között a beszédben (hang)átmenet képződik (Gay, 1978).

A koartikuláció jelenségét a jelen kísérletünk középpontjában álló *ija* és *ia* hangsorokon keresztül a következőképpen mutathatjuk be. Az [i] magánhangzó képzésekor a nyelv egyre magasabbra emelkedik, míg eléri a beszédhanghoz tartozónak vélt, bizonyos mértékben statikusan jelentkező artikulációs-akusztikai célt vagy beállítást. Az [i] célkonfigurációja megfigyelhető az 1. ábra F_1 -görbéin: a kérdéses felső nyelvállású magánhangzó az F_1 -görbéknek (azaz az *ija* és *ia* hangsoroknak az) elején, egy viszonylagosan alacsony frekvenciatartományban, statikusan valósul meg. Ezt követően a beszéd szerv-rendszer felkészül a következő beszédhang, jelen esetben az [ɒ] magánhangzó artikulációs-akusztikai céljának elérésére. Mivel az [ɒ] alsó nyelvállású szegmentum, a F_1 értéke az [i] után az 1. ábrán látható módon nő az *ija* és *ia* hangkapcsolatokban, az [ɒ] beszédhanghoz tartozó artikulációs-akusztikai cél pedig az F_1 -görbe pozitív kitéréseként figyelhető meg a hangsorozat végén. (Megjegyzendő, hogy az 1. ábrán látható módon a vizsgált vokális szakaszok kezdetén nagyon alacsony F_1 -értékek jelentkeztek, ami valószínűleg a megelőző nazális szegmentum koartikulációs hatásának tulajdonítható. Minden bizonnyal ennek köszönhető az is, hogy nem látunk az [i]-re egyébként jellemző negatív irányú kicsúcsosodást a görbéken.)

1. ábra. Az *ija* és *ia* vokalikus szakaszok F₁-görbéje a vokalikus szakasz normalizált időtartamára vetítve, a beszélő számára kényelmes („normál”) beszédtempóban (Deme & Juhász, 2024: 4. ábra nyomán) A kék téglalapok az F₁-lefutásának azon szakaszait jelölik, amelyek a statisztikai elemzés szerint szignifikánsan eltértek egymástól



A [j] approximáns esetében – a prototipikus magánhangzókban látottaktól eltérően – hiányozhat a jól kivehető artikulációs és akusztikus cél (Catford, 1988: 68), amit a formánsmenetekben gyakran egy kicsúcsosodás vagy akár hosszabb statikus szakasz, formánsplató (steady state, „tisztá fázis”) jelezne (Gay, 1978: 224). Különösen várható ez *i* mellett, mivel mind az [i] beszédhang, mind a [j] approximáns képzési helyüket tekintve az (alveolo)palatális régióban jönnek létre (Recasens, 2013). Ez a hiány, illetve a [j] hangátmenetszerű megjelenése figyelhető meg az 1. ábrán az [i] és az [v] célok között az írásban jelölt és nem jelölt [j] esetében egyaránt.

A fentiekkel együtt Juhász és Deme (2024) kísérletében alacsonyabb F₁-értékeket látunk, tehát magasabb nyelvállásfokot, illetve zártabb toldalékcsovet, tehát mássalhangzósbab ejtést sejtgetünk az *ija* hangsor két magánhangzós eleme közötti átmenetben, mint a helyesírásban *j*-t nem tartalmazó *ia* két magánhangzó(s cél)ja között. Ez pedig azt jelenti, hogy ezen korábbi eredmények szerint a beszélők spektrálisan elkülönítették azokat a *j* megvalósulásokat, amelyeket az írásképb megjelenít (pl. *ija*) azoktól, amiket nem (pl. *ia*).

Megjegyzendő még, hogy az idézett elemzésben a különbség csak a hangszín mentén jelentkezett, ugyanis a beszélők a valódi szavakban időtartamukban nem különítették el spektrálisan a *j*-t az írásképbben jelölő és nem jelölő formákat, szemben a korábban bemutatott álszavas kísérlet eredményeivel, ahol a *j*-t írásban is tükröztető formák ejtésében hosszabb vokalikus szakaszokat mértek a szerzők (Juhász & Deme, 2024). Ebből arra következtethetünk, hogy az álszavakat elemző kísérletben a *j*-t tartalmazó és nem tartalmazó írásképbű (ál)szavak akusztikai

elkülönítését az álszavakban várható eltúlzott artikuláció fokozhatta (vö. Maxwell et al., 2015).

A jelen elemzés a fentebb bemutatottak folytatása és kiterjesztése. Itt ismételtén a *j* graféma helyesírásban való megjelenésének és hiányának ejtésre gyakorolt hatását vizsgáljuk, ezúttal azonban abban a tekintetben, hogy a beszédsebesség, és kiemelten a beszédsebesség fokozása hogyan befolyásolja a kérdéses *ia* és *ija* szekvenciák közötti spektrális elkülönítést. Ez a kérdés azért releváns, mert megnövekedett beszédsebességben a hangsorok ejtésében törlés(ek)e)t, illetve redukciós jelenségeket várunk: kísérleti eredmények is alátámasztják, hogy felgyorsított beszédben a beszédhangok képzésére rendelkezésre álló rövidebb idő miatt a beszédképző szervek nem tudnak ugyanakkora utat megtenni, mint lassabb sebesség esetén, így nem érik el a beszédhangokra a feltételezések szerint jellemző, idealisztikus artikulációs-akusztikai célt (Lindblom, 1963). Ez pedig ahhoz vezethet, hogy a felgyorsított beszédben még a nyelvileg releváns hangzásbeli szembenállások kifejeződése is csökkenhet. Ugyan a nyelvileg releváns egyes kontrasztok teljes semlegesedését a magyar nyelvre eddig még nem mutatták ki – Deme és munkatársai (2024) például azt találták, hogy a hosszúsági magánhangzó-kontrasztok felgyorsított beszédben is megőrződnek bizonyos mértékben –, kérdésként merül fel, hogy mi történik a felgyorsított beszédben a jelen kísérletben elemzett hangsorokkal, ahol a feltételezések szerint pusztán az írásképek eltérései idézik elő az akusztikai szerkezetben tapasztalt különbségeket, tehát nem látszik indokoltnak a megőrzésük.

1.4. A jelen kutatás kérdése

Egy korábbi kísérletben magyar anyanyelvű beszélők izolált ejtésben spektrálisan elkülönítették a *zoknija* típusú szavakban megjelenő *ija* hangsort a *szoknia* típusú szavak *ia* hangsorától. Ez feltehetően az ortográfiai megjelenésnek volt köszönhető. A jelen kísérletben azt vizsgáljuk, hogy ez akusztikai elkülönítés mennyire áll ellen a beszédtempó gyorsításából következő redukciós folyamatoknak. Másként megfogalmazva azt elemezzük, hogy a kényelmes beszédsebességben tapasztalt eltérések megvalósulnak-e gyors beszédtempóban is.

Ha egy beszélőt egy felolvasási feladatban rövid, kétszavas megnyilatkozások felolvasására kérünk – mint a fentebb bemutatott, kényelmes beszédsebességhez kapcsolódó kísérlet esetében –, akkor arra számíthatunk, hogy az ejtés formálásában az ortográfiai megjelenítés hatása a meghatározó. Azonban akkor, ha ugyanazon megnyilatkozás ismételtetésével fokozzuk az ejtés automatizáltságát, illetve emellett még gyorsítottunk is a beszélővel a beszédének sebességén (mint ahogyan azt a jelen kísérlet esetében is tettük), arra számíthatunk, hogy az ejtés gazdaságosabbá válik, és redukció jelentkezik, a begyakorlottság foka nő, az ejtésben jelentkező (vélt) ortográfiai hatás pedig csökken (Jacobs et al., 2015; Greisbach, 1992; Juhász, 2024b; Deme et al., 2024).

Ennek mentén arra számítunk, hogy az *ija* és *ia* szekvenciák kényelmes (illetve „normál”, vö. lentebb) beszédsebességben megfigyelt spektrális elkülönítése a felgyorsított beszédben nem valósul meg. Az akusztikai mérőszámokat tekintve ez azt jelenti, hogy az *ija* és *ia* vokalikus szakaszok időtartama között – a normál tempójú ejtéshez hasonlóan – nem várunk eltérést felgyorsított beszédben sem. A spektrális minőséget tekintve pedig arra számítunk, hogy az *ija* és *ia* szekvenciák között a normál tempóban megfigyelt eltérés a beszédsebesség növelésének hatására eltűnik.

2. Módszertan

A kísérletben 15 felnőtt magyar anyanyelvű nő vett részt, akikkel hangfelvételeket készítettünk. A feladatuk az volt, hogy a *Legyen* <célszó>! hordozómondatban olvassanak fel olyan szavakat, amelyekben a helyesírás jelöli a *j*-t (*Andija*, *zoknija*, *haknija*, *pónija*) vagy nem jelöli azt (*adnia*, *szoknia*, *raknia*, *rónia*). (A példákban félkövérrel jelöljük a hangsoroknak a jelen vizsgálatban elemzett vokalikus szakaszait, a példákat Rákosi & Laczkó, 2008: 166, illetve Siptár, 2011: 153 nyomán alkottuk meg.) A célszavak a két csoportban nem képeztek egymással minimális párokat, így nem számíthatunk eltúlzott ejtésre annak következtében, hogy a beszélő – felismerve a szópárok közötti hasonlóságot – túlságosan igyekezne az egyébként alig vagy nem megjelenő különbségeket érzékeltetni. Ugyanakkor fonetikai szempontból jól összevethetők a vizsgált párok, hiszen a szavak magánhangzószerű szakaszai a megelőző mássalhangzó és a mássalhangzón túli magánhangzó minősége szerint is illesztettek (az *Andija* szó kivételével minden esetben [n] áll a kérdéses szakasz előtt, de ebben a szóban is azonos a zöngéesség és a képzés helye a további szavak megfelelő mássalhangzójával), továbbá a szótagszerkezet és a szótagszámok páronként egyeznek. A hordozómondat azt biztosította, hogy a célszavak prozódiai tekintetben egyfelől azonos, másfelől kiemelt (hangsúlyos és frázisvégi) helyzetben legyenek. A hangfelvételeket a beszédsebesség szerint két feltételben rögzítettük: a) egy, a számunkra kényelmes beszédsebességben (erre egyszerűsítve „normál” beszédként hivatkozunk az elemzésben), és b) maximálisan felgyorsított beszédsebességben (erre „gyors” beszédként hivatkozunk a továbbiakban). A maximálisan felgyorsított beszédet Greisbach (1992) módszerével értük el. A beszélők minden hordozómondatot (és így célszót) többszöri ismétlésben ejtettek. Az első ejtés esetében csak annyi instrukciót kaptak, hogy ejtsék ki az adott mondatot kényelmes tempóban; ezt az ejtést tekintettük „normál” beszédnek az elemzésben. Ez után arra kértük a résztvevőket, hogy kezdjék el ismételtetni az adott hordozómondatot úgy, hogy minden ejtéssel igyekezzenek egyre gyorsabban és gyorsabban beszélni addig, amíg ki nem fogynak a levegőből vagy meg nem botlik a nyelvük. Minden résztvevő 6 ilyen gyorsuló sorozatot állított elő (azaz egy normál sebességű változatot, és ezt követően gyorsabb ejtésű változatokat), amihez randomizált sorrendben vetítettük

a felolvasandó megnyilatkozásokat egy számítógép monitorjára. Ezekből az ismétlésekből a legrövidebbet (tehát a leggyorsabb beszédsebességűt) tekintettük „gyors” beszédnek. Így összesen ($15 \text{ kísérleti személy} \times 8 \text{ hangsor} \times 6 \text{ ismétlés} \times 2 \text{ beszédsebességi feltétel} =$) 1480 hangsort vizsgáltunk.

A felvételek az ELTE Alkalmazott Nyelvészeti és Fonetikai Tanszékének laboratóriumában készültek, csendesített szobában. A hangfelvételeket az Audacity (The Audacity Team, 2019) programmal készítettük egy külső hangkártya és egy omnidirekcionális kondenzátoros fejmikrofon segítségével.

A hangfelvételeket akusztikai elemzésnek vetettük alá a Praat programban (Boersma & Weenink, 2020). Kézzel címkéztük a vizsgált hangsorok vokalikus szakaszát. Mint azt már említettük, a hangsorok első, izolált ejtésű megvalósulását tekintettük normál tempójú megvalósulásnak (ehhez kapcsolódóan mutattuk be az eredményeket a korábbi tanulmányunkban, vö. Deme & Juhász, 2024, ill. a jelen tanulmányban az 1. ábrán). Az ismétlések közül a vokalikus szakaszok időtartamát összevetve választottuk ki a legrövidebb időtartamú realizációt, és ezt tekintettük a gyors tempójú megvalósulásnak. A vokalikus szakaszok időtartamát, illetve az F_1 formáns frekvenciáját mértük automatikusan, 5 ms-onként a Burg algoritmus segítségével. Az F_1 idői lefutásának elemzésével azt kívántuk megragadni, hogy hogyan alakul a hangsorok vokalikus részében a nyelvemelkedés foka, illetve a létrejött szűkület keresztmetszete.

A vokalikus szakaszok időtartamát lineáris kevert modellekkel (Kuznetsova et al., 2017) elemeztük az R programban (R Core Team, 2020). Ezekben a modellekben a független változó a *j típusa* (helyesírásban megjelenő/nem megjelenő), illetve a *beszédsebesség* (normál/gyors) interakciója volt, random hatásként a *beszélőt* adtuk meg. Az ide kapcsolódó ábrákat a ggplot2 (Wickham, 2016) csomag segítségével készítettük.

Az F_1 -értékek idői lefutását dinamikus elemzésben, általánosított additív kevert modellekkel (generalised additive mixed model, GAMM) elemeztük az mgcv csomag (Wood, 2017) segítségével. A GAMM-ok nemlineáris adatsorok elemzésére szolgálnak, ezért a formánsmenetek esetében – ahol görbék lefutását elemezzük – jobb illesztést biztosítanak a lineáris modelleknél (Wieling, 2018). Az alapmodellben az F_1 frekvenciáját mint függő változót vizsgáltuk a *normalizált időtartam* függvényében. (A vokalikus szakasz normalizált időtartamát úgy kaptuk meg, hogy az F_1 mérési pontjaiként szolgáló, 5 ms-onként növekvő időértékeket arányítottuk a teljes vokalikus szakasz időtartamával.) Emellett a modellt egy parametrikus faktor változóval is kiegészítettük. Ez a változó négy szinttel rendelkezett, amely szintek a *j* ortográfiai megjelenésének és a beszédsebesség szerinti feltételeknek a kombinációit kódolták (*ia/ija* és normál/gyors tempó). A modellbe emellett random simítást illesztettünk tokenenként, majd az autokorreláció ellenőrzése (acf.resid()) után annak korrekcióját is elvégeztük. Végül az így kapott modellt a gam.check() paranccsal ellenőriztük. A modellek által becsült F_1 -görbék minden esetben 95%-os

konfidencia-intervallummal ábrázoltuk a ggplot2 csomag segítségével (Wickham 2016).

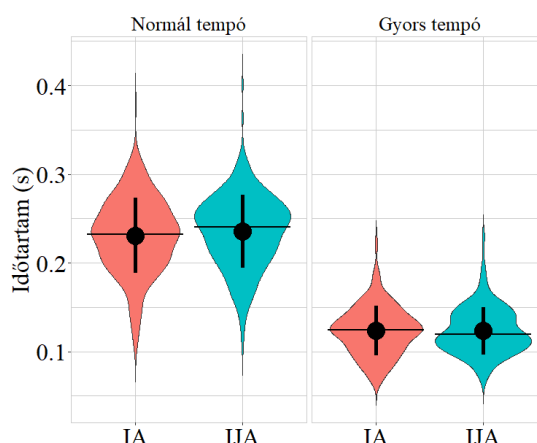
3. Eredmények

Az alábbiakban előbb az időtartamok, majd az F_1 frekvenciaértékére kapott adatok elemzését mutatjuk be.

3.1. A célszavak vokalikus szakaszainak időtartama normál és gyors beszédtempóban

A 2. ábra a célszavakban megvalósult vokalikus szakaszok időtartamát mutatja be normál, illetve gyors beszédtempóban. A kevert modell eredményei szerint a vokalikus szakaszok időtartamára kizárólag a *beszédtempó* független változó gyakorolt szignifikáns hatást ($F(1,1415) = 4357, p < 0,001$), míg a *j-k* típusa nem. A post hoc teszt eredményei szerint a különböző beszédtempóban megvalósult szavak vokalikus szakaszainak időtartama eltért egymástól ($p < 0,001$), azonban a *j* helyesírásbeli jelöltsége szerint nem mutatkozott eltérés sem a normál, sem a gyors beszédtempón belül (2. ábra). Habár a különbségek mindössze a tendencia szintjén jelentek meg, megemlítendő, hogy míg normál tempóban a vokalikus szakaszok esetében az *ija* szekvencia rendelkezett valamivel magasabb átlaggal és mediánnal ($M_{\text{normál ija}} = 235,4$ ms, $Mdn_{\text{normál ija}} = 239,8$ ms) az *ia*-hoz képest ($M_{\text{normál ia}} = 231,1$ ms, $Mdn_{\text{normál ia}} = 232,8$ ms), addig ez a mintázat a gyors tempójú ejtésben éppen ezzel ellentétes volt ($M_{\text{gyors ija}} = 123,2$ ms; $Mdn_{\text{gyors ija}} = 119,4$ ms; $M_{\text{gyors ia}} = 123,4$ ms; $Mdn_{\text{gyors ia}} = 124,4$ ms).

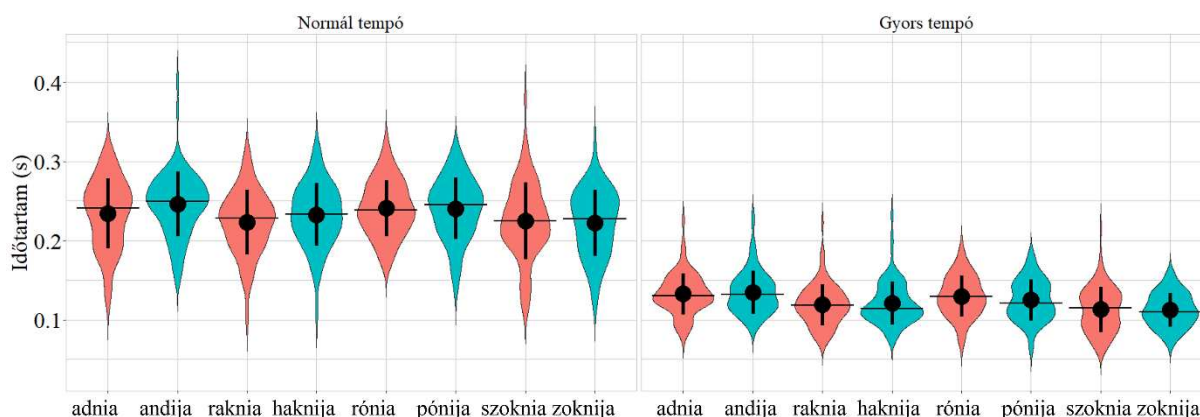
2. ábra. Az *ija* és *ia* hangsorok vokalikus szakaszainak időtartama normál, illetve gyors beszédtempóban a *j* helyesírásbeli megjelenésének függvényében (a fekete pont az átlagot, a vízszintes vonal a mediánt, a függőleges vonal pedig a szórást jeleníti meg)



Ha a vokalikus szakaszok időtartamát a célszavak szerinti bontásban vizsgáljuk meg, akkor általánosságban a fentihez hasonló tendenciát láthatunk: a gyors tempójú ejtésben egyetlen szópáron kívül medián értékét tekintve minden esetben az *ija* szekvencia időtartama bizonyult alacsonyabbnak a az *ia*-hoz képest. Ez alól

a mintázat alól az *adnia-Andija* szópair számít kivételnek, ahol éppen ezzel ellentétes tendenciát találtunk (3. ábra).

3. ábra. A különböző szavakhoz tartozó *ia* és *ija* hangsorok időtartama normál, illetve gyors beszédtempóban a hangsort tartalmazó szavak szerinti bontásban (a fekete pont az átlagot, a vízszintes vonal a mediánt, a függőleges vonal pedig a szórást jeleníti meg)

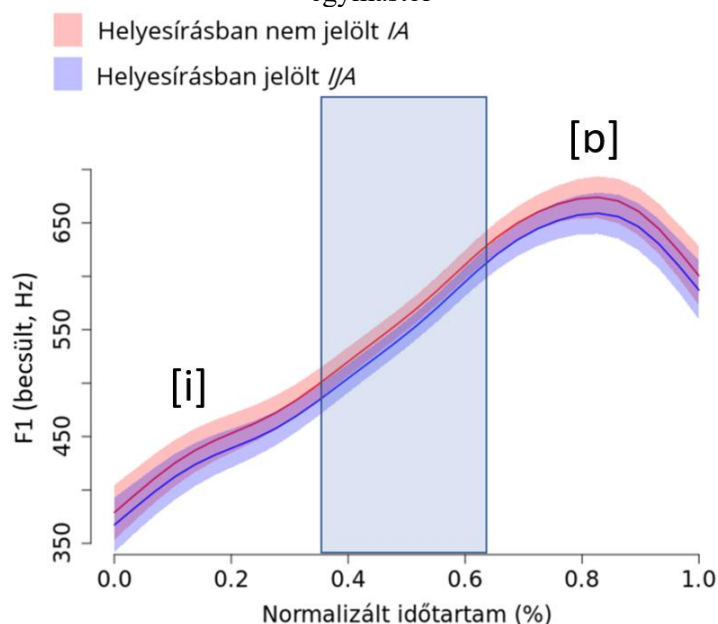


3.2. A vokális szakaszokban mért F_1 idői lefutása

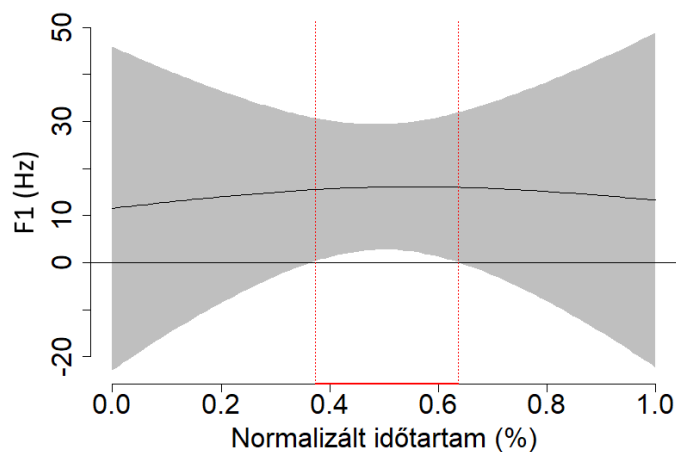
A gyors beszédtempóban ejtett *ia* és *ija* hangsorokban mért frekvenciaértékek mind az F_1 -görbék becsült átlagában, mind lefutásukban szignifikáns eltérést mutattak (az *ia* esetében a parametrikus eredmények *becsült átlaga* = 551,6 Hz, $p < 0,001$, smooth együttthatók: $edf = 8,0$, $F = 44,4$; míg az *ija* esetében a parametrikus együttthatók *becsült átlaga* = 536,9 Hz, $p < 0,001$, smooth együttthatók: $edf = 1,27$; $F = 0,1$).

Az F_1 -görbék GAMM által becsült megvalósulását a 4. ábra, az F_1 -görbék különbségét az 5. ábra illusztrálja, és mindkét ábrán megfigyelhető a két görbe szignifikánsan eltérő szakasza. A 4. ábrán – hasonlóan az 1. ábrán bemutatottakhoz – az F_1 -értékeknek a vokális szakasz normalizált időtartamán belüli változása látható: a vokális szakasz elején a többségében nazális beszédhangkörnyezetből fakadóan az F_1 -görbe relatíve alacsony frekvenciasávból indul, ezután egy minimális kitérést, vagyis az [i] akusztikai céljának elérését követően egy lineáris átmenettel közelíti meg az [v]-ra jellemző, relatíve magas F_1 -értéket. Ezen az átmeneten belül találjuk az *ija* és az *ia* F_1 -görbéinek szignifikánsan eltérő szakaszát a normalizált időtartam 37% és 63%-a között. A szignifikánsan eltérő szakasztól eltekintve az mondható el, hogy az *ija* F_1 -görbéje a tendencia szintjén a normalizált időtartam egészében az *ia* megvalósulásához képest alacsonyabb F_1 -értékekkel realizálódott (4. ábra).

4. ábra. Az *ija* és *ia* vokális szakaszok F_1 -menete a gyors tempójú ejtésében. A görbék fölött megjelenő [i], illetve [ɒ] a hangsorban megvalósuló akusztikai céljait jelöli, a kék téglalapok az F_1 -lefutásának azon szakaszait jelölik, amelyek a statisztikai elemzés szerint szignifikánsan eltértek egymástól



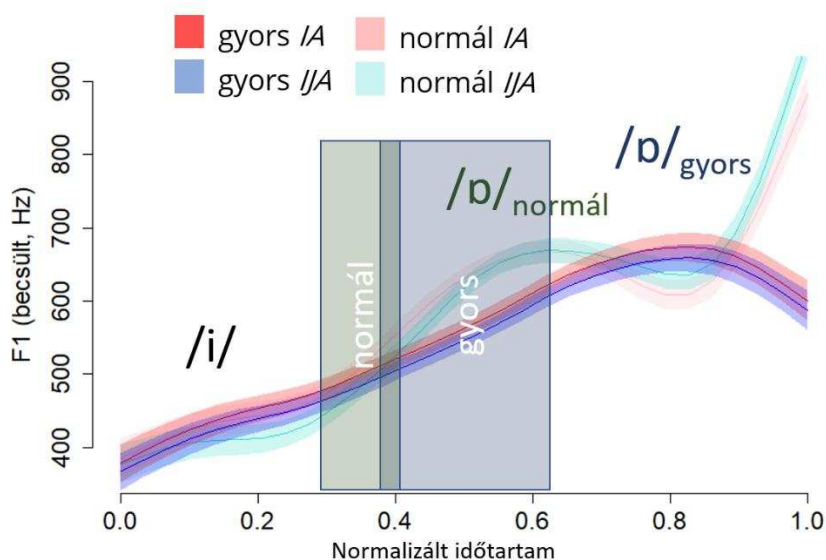
5. ábra. A célszavakban ejtett *ija* és *ia* hangsorok F_1 -görbéinek különbsége gyors tempóban. A görbék szignifikánsan eltérő szakaszát a piros körvonal jelzi



Összehasonlítva az *ia* és *ija* vokális szekvenciák megvalósulását normál és gyors beszédtempóban azt láthatjuk, hogy a beszélők a két beszédtempóban jelentősen eltérő F_1 -görbéket produkáltak (6. ábra). Mivel az F_1 -görbéket minden esetben a normalizált időtartamra vetítve jelenítettük meg, ezért egyrésről elmondható, hogy a közel fele olyan hosszú időtartamú gyors tempójú görbéket „széthúztuk” a normál időtartamú megvalósulásoknak megfelelően. De nem ez az a tényező, ami alapvetően meghatározza az akusztikai célok helyzetét a normalizált időtartamon belül, sokkal inkább a koartikuláció hatása és a vokális szekvenciák pozíciója a megnyilatkozásokban.

A normál tempójú megvalósulásokat izolált ejtésben rögzítettük, tehát szünet következett utánuk, ahol a frázis végén a zöngékezés fokozatos elhalása volt megfigyelhető. Ez pedig a normalizált időtartamon belül a normál tempójú ejtésben a hangsor eleje felé tolta az akusztikai célok pozícióját. Ezzel szemben a gyors tempójú ejtésben a vizsgált vokális szakaszt rögtön a célmondat következő ismétlése követte szünet nélkül, vagyis a *Legyen...* szó [l] beszédhangja – amelyet természetesen már nem szegmentáltunk a vokális szakasz részeként. Az F₁-görbe a [l] (és az azt követő hangok) koartikulációs hatásából fakadóan nem kanyarodik felfelé (a normál tempójú realizációjától eltérően). Az F₁-görbék általános formáját illetően megfigyelhető, hogy a normál tempójú ejtésben az [i] magánhangzó akusztikai célját követő statikusabb eltűnt a gyors tempójú ejtésben, és a helyét egy kvázi-lineáris átmenet vette át. Ezzel szemben az [ɒ] beszédhang akusztikai célját nem befolyásolta szembetűnően a beszédtempó, hiszen mind normál, mind gyors tempóban hasonló értékekkel realizálódott (6. ábra).

6. ábra. A normál és gyors tempóban ejtett *ia* és *ija* hangsorok F₁-görbéi. A színes téglalapok az F₁-lefutásának azon szakaszait jelölik, amelyekben a statisztikai elemzés szerint az *ia* és *ija* szignifikánsan eltértek egymástól az egyes sebességfeltételeken belül



4. Következtetések

A kísérletben arra a kérdésre kerestük a választ, hogy hogyan befolyásolja a gyors beszédtempó az ortográfiailag jelölt, illetve az ortográfiában nem jelölt *j*-realizációk produkcióját. Ezt a kérdést toldalékolt főnevek (pl. *zoknija*), illetve ragozott főnévi igenevek (pl. *szoknia*) ejtésében megvalósuló *ija* és *ia* hangsorok összevetésével vizsgáltunk dinamikus módszerekkel.

Korábbi eredmények szerint a magyar beszélők kényelmes beszédsebességben (erre az egyszerűség kedvéért a jelen tanulmányban „normál” sebességként

hivatkoztunk) az időtartamuk szempontjából nem, azonban spektrális tulajdonságaik mentén elkülönítették a kérdéses fentebb említett *ija* és *ia* hangsorokat. (Deme & Juhász, 2024). A spektrális különbség normál tempóban a toldalékcso zárttságával összefüggésben álló F_1 -érték mentén jelentkezett: a magyar beszélők a célszavakban megjelenő *ija* hangsorban az [i] és [ɒ] akusztikai célok közötti átmenetet, vagyis a *j* approximánst az *ia* szakaszban látottakhoz képest alacsonyabb F_1 -értékekkel, azaz – vélhetően – zártabb toldalékcsoval, mássalhangzóssalban produkálták. A jelen kísérletben azt vizsgáltuk meg, hogy ez az ortográfiailag motivált, vagyis a *j* graféma megjelenésének vagy hiányának köszönhető eltérés a hangsorok között mennyire ellenálló a beszédsebesség felgyorsításával járó redukcióval szemben. Abból indultunk ki, hogy míg a kényelmes sebességű ejtésben az ortográfiai hatás meghatározó szereppel bírhat, és készítheti a beszélőket valamifajta eltérés létrehozására a kérdéses hangsorok között, addig a kísérleti beszédfeladatban kért többszöri ismétlés fokozza az ejtés begyakorlottságát, a beszédsebesség növelése pedig az ejtés gazdaságosságát, ami redukciós folyamatok érvényesüléséhez vezet. Ennélfogva azt vártuk, hogy az *ija* és az *ia* a felgyorsított beszédben már nem különülnek el egymástól.

A kísérlet eredményei megerősítették azt a feltételezést, hogy a célszavakban ejtett *ija* és *ia* hangsorok a normál tempójú ejtéshez hasonlóan a gyors beszédtempóban sem térnek el egymástól időtartamukban: a beszédsebesség gyorsításának hatására a helyesírás által jelölt és nem jelölt *j*-t tartalmazó hangsorok ugyanolyan mértékben redukálódtak az időtartam szempontjából. Ez közvetetten utalhat arra, hogy a beszélők azonos számú elemet (beszédhangot) ejtettek ezekben a hangsorokban.

Azt is feltételeztük, hogy a gyors tempójú ejtésben – amit a megnyilatkozások sokszoros ismétlése hív elő – az ejtés automatizációjából és az időtartambeli megszorításokból fakadóan az *ija* és az *ia* szakaszok közötti normál beszédtempóban megfigyelt spektrális eltérés is eltűnik. Ezt azonban a kísérlet eredményei nem támasztották alá: az *ija* és *ia* hangsorok közötti, a toldalékcso zártságát, illetve a nyelvemelkedés fokát számszerűsítő F_1 értéke mentén a kényelmes beszédsebességben tapasztalt eltérés gyors beszédtempóban is jelentkezett. Megjegyzendő, hogy az F_1 -görbék eltérő szakasza a felgyorsított beszédben a normál beszédben látott időablakhoz képest a normalizált időtartam egy későbbi pontján mutatkozott: míg normál tempójú ejtésben ez a szakasz az időtartam 28% és 40%-a között, addig a gyors tempójú ejtés esetében az időtartam 37% és 63%-a között valósult meg. Ahogyan azonban azt már fentebb is említettük, ez fakadhat pusztán koartikulációs hatásokból, és lehet a célszekvenciát követő [l] mássalhangzó (és a további hangsorok) koartikulációs hatásának eredménye. Az F_1 -görbék különbségének eltérő időzítése tehát alapvetően nem mond ellent a levonható következtetéseknek: az *ija* hangsorban az [i]-hez és az [ɒ]-hoz tartozó akusztikai célok közötti átmenet mind a normál, mind a gyors tempójú ejtésben alacsonyabb F_1 -értékkel, tehát zártabb

toldalékcsovéval, mássalhangzóssabban valósult meg, mint az *ia* szekvenciában. Ez pedig azt jelenti, hogy a beszélők még az automatizált, gyors tempójú produkcióban is megőrizték az *ija* és *ia* elkülönítését az ejtésükben. A fenti eredmény arra utal, hogy a helyesírás hatása, tehát az a tény, hogy a helyesírás megjelenít vagy nem jelenít meg egy *j* elemet a kérdéses szavakban olyan mértékben formálja a megvalósítást, hogy a beszélők még az artikulációs szervek felgyorsított működése és az időtartambeli megkötöttség ellenére is különbséget produkálnak a szóban forgó fonológiai nem kontrasztív *ija* és *ia* hangsorok között.

A gyors beszédtempó nem csak a szignifikánsan eltérő szakaszában volt hatással az F_1 -görbék realizációjára: érdemes megemlíteni az [i] megvalósulásával kapcsolatos eltérést a normál és gyors beszédtempók között. Normál beszédtempóban az [i] akusztikai célja – egy prototipikus ejtésű magánhangzóra jellemzően – egy viszonylagosan statikus fázissal valósult meg a formánsmenetben, ahol az F_1 értéke az adott intervallumban megközelítőleg állandónak volt tekinthető. Ezzel szemben gyors tempóban az [i] statikusan megvalósuló akusztikai céljának helyét egy kvázi-lineáris tranzíció vette át, tehát az [i] tulajdonképpen a leggazdaságosabb átmenetként valósult meg a megelőző nazális beszédhang és az [v] akusztikai célja között. Ez az eltérés a normál és gyors tempójú megvalósulások között abba az irányba mutat, hogy a gyors beszédtempó hatására az [i] a teljes értékű magánhangzókra jellemző „komplexebb”, vagyis több fázisú megvalósulása helyett egyfajta siklóhanggá (approximánssá) redukálódott. Ráadásul az [i] itt leírt redukciója mind az *ija*, mind az *ia* hangsorokban végbement. Az [i] magánhangzó redukálódása olyan szempontból nem meglepő, hogy például az *ia* magánhangzó-kapcsolat esetében fennálló hiátushelyzetnek fonológiai szempontból a *j* betoldása mellett egy lehetséges feloldása lehet az egyik magánhangzó (jelen esetben az [i]) siklóhanggá alakítása, és ilyenén módon a két szótagú konstrukció redukálása egy szótagúvá (Siptár, 2011). Úgy látszik tehát, hogy ennek a stratégiának a megjelenését fokozhatja a beszédsebesség gyorsítása.

A kísérletben az [v] akusztikai célja jelentősen ellenállóbbnak mutatkozott a gyors beszédtempóval szemben, mint az [i]-é: a normál és gyors tempójú F_1 -görbék megvalósulását összevetve azt állíthatjuk, hogy az [v] akusztikai célja az F_1 tekintetében megközelítőleg 650 Hz volt beszédtempótól függetlenül, azaz – legalábbis az akusztikai szerkezetnek az itt vizsgált dimenziójában – nem jelentkezett az [v] ejtésének centralizációja. Ami azonban egy kisebb mértékű, de mégis szembevető eltérés az [v] beszédhang normál és gyors tempójú realizációja között, az az, hogy míg a normál tempójú ejtés esetében az *ija* és az *ia* hangsorokban az [v]-k akusztikai céljához tartozó F_1 -görbék teljesen átfedtek egymással, addig gyors tempóban nem, hiszen az *ija* szekvenciában látott, a tendencia szintjén általánosan alacsonyabb F_1 -értékek az [v] megvalósulásában is jelentkeztek. Ez azt jelenti, hogy az *ija* gyors tempójú ejtésében tapasztalt *j*-

megvalósulás előrefelé ható – tendenciaszintű – koartikulációs hatást gyakorolt az [p] megvalósulására is.

Összegezve: a kísérlet eredményei rámutattak arra, hogy a gyors beszédtempó hatására sem tűnik el a kényelmes sebességű beszédben tapasztalt spektrális elkülönítés a ragozott főnévi igenevek (pl. *szoknia*) *ia* szekvenciája, valamint a toldalékolt főnevek (pl. *zoknija*) *ija* szekvenciája között, ami közvetetten megerősíti azt az elképzelést, hogy az ortográfia, vagyis az íráskép hatással van a hangsorok ejtésére. Az eredmények szerint ugyanis mind a kényelmes, mind a felgyorsított beszédben zártabb toldalékcsovéval ejtett, azaz mássalhangzóssabb a hangsorok [i] és [p] magánhangzója közötti átmenet akkor, ha a helyesírásban megjelenik a *j*. Mivel a *j* helyesírásbeli jelöltsége a hiátustöltő és fonemikus *j* körüli tudományos polémiahoz is kapcsolódik, vizsgálatunk általánosságban is hozzájárul a magyar *j*-realizációk akusztikai és ejtési tulajdonságainak mélyebb megértéséhez.

Köszönetnyilvánítás

Köszönjük a kutatásunkhoz kapcsolódó hasznos beszélgetéseket és észrevételeket Rebrus Péternek és Fejes Lászlónak, a jelen tanulmányhoz fűzött megjegyzéseket anonim lektorainknak, valamint köszönetünket fejezzük ki Markó Alexandrának, aki a tanulmány elkészítésében segítségünkre volt. A kutatás a Kulturális és Innovációs Minisztérium EKÖP-24 kódszámú Egyetemi Kiválósági Ösztöndíj Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával, a Bolyai János Kutatási Ösztöndíj szakmai támogatásával, valamint a TKA–DAAD 177375. sz. projektjének, az NKFIH FK128814 sz. projektjének támogatásával készült.

Irodalom

- Bassetti, B. (2008). Orthographic input and second language phonology. In T. Piske, T. & M. YoungScholten (eds.), *Input Matters in SLA* (191–206). Clevedon, UK: Multilingual Matters.
- Boersma, P. & Weenink, D. (2020). Praat: doing phonetics by computer [Computer program]. Version 6.1.15. (A letöltés ideje: 2019. november 4.).
- Cruttenden, A. (2008). *Gimson's pronunciation of English*. 7th edn. London: Hodder Education.
- Davidson, L. & Erker, D. (2014). Hiatus resolution in American English: The case against glide insertion. *Language*, 90(1), 482–514. <https://doi.org/10.1353/lan.2014.0028>
- Deme A. (2016). *A magánhangzók ejtése és észlelése a szopránénekben*. Budapest: ELTE Eötvös Kiadó.
- Deme A. & Juhász K. (2024). A helyesírásban jelölt és nem jelölt *j* ejtésének mássalhangzóssága az ejtés akusztikai vetülete alapján. *Alkalmazott Nyelvtudomány*, 2, 1–20.
- Deme, A., Juhász, K., Szánthó, Zs., Zsoldos, Sz., & Greisbach, R. (2024). German and Hungarian long and short vowels in fast speech. In *Proceedings of 5th International Symposium on Applied Phonetics (ISAPh 2024)* (5–10). Tartu: ISAPh.
- Fant, G. (1960). *Acoustic theory of speech production*. Mouton: The Hague.
- Gay, T. (1978). Effect of speaking rate on vowel formant movements. *The Journal of the Acoustical Society of America*, 63, 223–230. <https://doi.org/10.1121/1.381717>.
- Gósy M. (2014a). A palatális közelítőhang kétféle funkcióban. *Beszédkutatás*, 22, 17–40.
- Gósy, M. (2014b). Phonetic Properties of Hungarian Hiatus Fillers versus Intervocalic Palatal Approximants: Physical Realities and Theoretical Implications.
- Gósy M. (2018). Hiátusos kontextusok temporális mintázata gyermekek beszédében. *Beszédkutatás*, 2018, 110–133. 10.15775/Beszkut.2018.110-133.

- Greisbach, R. (1992). Reading aloud at maximal speed. *Speech Communication*, 11, (4–5), 469–473, [https://doi.org/10.1016/0167-6393\(92\)90053-A](https://doi.org/10.1016/0167-6393(92)90053-A).
- Heselwood, B. (2006). Final schwa and r-sandhi in RP English. *Leeds Working Papers in Linguistics and Phonetics*, 78–95.
- Hunt, E. H. (2009). *Acoustic characterization of the glides /j/ and /w/ in American English*. Cambridge, MA: MIT Doctoral Dissertation.
- Huszthy B. (2023). Az írás hatásai a kiejtés változására. In A. Cser & A. L. Nemesi (szerk.), *Nyelvelmélet és diakronia 5.* (307–332). Budapest: Pázmány Péter Katolikus Egyetem, Bölcsészeti és Társadalomtudományi Kar.
- Jacobs, C. L., Yiu, L. K., Watson, D. G. & Dell, G. S. (2015). Why are repeated words produced with reduced durations? Evidence from inner speech and homophone production. *Journal of Memory and Language*, 84, 37–48.
- Jagers, Z. (2018). Evidence and characterization of a glide-vowel distinction in American English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 9, 1–27. <https://doi.org/10.5334/labphon.36>.
- Juhász K. (2024a). *A mandarin beszédhangok produkciója kínaiul tanuló magyar anyanyelvűek ejtésében*. PhD disszertáció. Budapest: Eötvös Loránd Tudományegyetem. <https://doi.org/10.15476/ELTE.2024.067>
- Juhász K. (2024b). Az ismétlések hatására bekövetkező f0-redukció kérdése a magyar egy szótagú intonációs kontúrok megvalósulásában: az egyéni produkciók vizsgálata. *Alkalmazott Nyelvtudomány Különszám*, 2, 2–23. <https://doi.org/10.18460/ANY.K.2024.2.001>.
- Juhász K. & Deme A. (2022). Palatális approximánsok a kínaiban és a magyarban – a kínai alveolopalatális /ɕ/ szibilánst követő vokális szakasz produkciója kínaiul tanuló magyar anyanyelvűeknél. *Általános Nyelvészeti Tanulmányok*, 34, 287–332.
- Juhász K. & Deme A. (2024). Az *iá* és az *ijá* vokalikus hangsorok megvalósulása magyar álszavakban a nyelvváltás akusztikai vetületének szempontjából. *Beszéd Tudomány*, 9–41.
- Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Ladefoged, P. (1975). *A Course in Phonetics*. New York: Harcourt.
- Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35, 1773–1781.
- Maxwell, O. Baker, B. Bundgaard-Nielsen, R. L. & Fletcher, J. (2015). A comparison of the acoustics of nonsense and real word stimuli: coronal stops in Bengali. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, 10–14 August 2015, Glasgow, Scotland, UK.
- Menyhárt K. (2006). Koartikulációs folyamatok két magánhangzó kapcsolatában. *Beszéd kutatás*, 2006, 44–56.
- Olaszy G. (2010). Az [i] + V és V + [i] hangkapcsolódások akusztikai elemzése a hiátustöltés magyarázatához. *Beszéd kutatás*, 2010, 76–81.
- R Core Team (2020). *R Foundation for Statistical Computing*, Elérhető: <https://www.R-project.org/>. 3.6.1-es verzió. (A letöltés ideje: 2020. április 12.).
- Rákosi, Gy. & Laczkó, T. (2008). On the categorial status of agreement-marked infinitives in Hungarian. In Piñon, Christofer & Szentgyörgyi Szilárd (eds.) *Approaches to Hungarian vol. 10, Papers from the Veszprém Conference* (145–172). Budapest: Akadémiai Kiadó.
- Rebrus P. & Kálmán L. (2009). Valóban megmagyarázhatatlanok a magyar infinitívusz toldalékai? In Maleczki M. & Németh T. E. (szerk.), *A mai magyar nyelv leírásának újabb módszerei VII. Modellek, elméletek és elvek érvényessége nyelvi adatok tükrében* (153–156). Szeged: Szegedi Tudományegyetem Általános Nyelvészeti és Magyar Nyelvészeti Tanszéke.
- Recasens, D. (2013). On the articulatory classification of (alveolo)palatal consonants. *Journal of the International Phonetic Association*, 43(1), 1–22.
- Siptár P. (2002). Hiátus. In Hunyadi László (szerk.), *Kísérleti fonetika, laboratóriumi fonológia* (85–98). Debrecen: Debreceni Egyetem Kossuth Egyetemi Kiadója.
- Siptár P. (2005). A magánhangzó-kapcsolatok fonológiájából. *Magyar Nyelv*, 101, 282–304.
- Siptár P. (2011). Alakváltozatok, allomorfok, alternációk. *Magyar Nyelv*, 107, 147–160.

- Siptár P. (2013). Palatálisok. In A. Benő, E. Fazakas, & E. Kádár (szerk.), „...hogy legyen a víznek lefolyása...”: *Köszöntő kötet Szilágyi N. Sándor tiszteletére* (433–448). Kolozsvár: Erdélyi Múzeum Egyesület. Letöltés: <http://archive.nytud.hu/oszt/elmnyelv/siptar/publ/palatalisok.pdf>.
- Siptár, P. & Törkenczy, M. (2000/2007). *The Phonology of Hungarian*. New York: Oxford University Press.
- Stevens, K. (2000). *Acoustic Phonetics*. Massachusetts: MIT Press. <https://doi.org/10.7551/mitpress/1072.001.0001>
- The Audacity Team (2019). *Audacity*. 2.4.2-es verzió. Elérhető: <http://audacityteam.org/> (A letöltés ideje: 2020. március 13.).
- van Heuven, V. & Hoos, A. (1991). Hiatus deletion, phonological rule or phonetic coarticulation? *Linguistics in the Netherlands, 1991*, 61–70.
- van Rij, J., Wieling, M., Baayen, H. R. & van Rijn, D. (2020). *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs*. 4.1.2-es R csomag-verzió. (A letöltés ideje: 2021. november 1.)
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.
- Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, 70, 86–116. <https://doi.org/10.1016/j.wocn.2018.03.002>
- Wood, S. (2017). *Generalized Additive Models: An Introduction with R*. Boca Raton: Chapman and Hall. <https://doi.org/10.1201/9781315370279>

GÓSY MÁRIA¹ – VLADÁR ZSUZSA²

¹HUN-REN Nyelvtudományi Kutatóközpont és Eötvös Loránd Tudományegyetem
gosy.maria@nytud.hun-ren.hu
<https://orcid.org/0000-0003-4336-3007>

²Nemzeti Közszolgálati Egyetem
Constantinovitsne.Vladar.Zsuzsanna@uni-nke.hu
<https://orcid.org/0000-0001-6356-3407>

Gósy Mária – Vladár Zsuzsa: Az *izé* eredete, története és mai használata
Alkalmazott Nyelvtudomány, XXV. évfolyam, 2025/1. szám, 44–66.
[doi:http://dx.doi.org/10.18460/ANY.2025.1.003](http://dx.doi.org/10.18460/ANY.2025.1.003)

Az *izé* eredete, története és mai használata

In Hungarian, there is an item of the language's vocabulary, *izé*, that can be found probably in each Hungarian speaker's mental lexicon. Although *izé* has no stable semantic content and practically no meaning, it can refer to any lexeme, substitute any part of speech, and take various suffixes. In addition, there is another function of this word, which is to help the speaker gain time in the case of a speech planning trouble at the level of thoughts, grammar, or finding the intended word. According to its function, we may identify it as a filler; however, fillers of the analyzed languages are real words with unambiguous semantic content. In fact, *izé* functions as a specific filler in Hungarian speech, providing extra time for the speaker to overcome a speech planning trouble, but it is not a filled pause, and it is entirely different from diverse discourse markers.

The sound sequence *izé* can take on diverse syntactic roles, such as verb, noun, adjective (etc.), with the proper forms in grammatical structures since it can be conjugated and can take on a number of declination suffixes. While producing *izé*, the speaker may find the desired word to utter (e.g., *így böködté a izének a vakondnak a fejét* 'he like prodded the head of the *izé* the mole'), or may continue along with the next thought (*menjünk be majd oda izé milyen jó lesz ott* 'let's go in there *izé* there it will be very good').

For a historical analysis, we searched for occurrences of *izé* in dictionaries, grammars, scientific papers, various written materials, transcribed texts of verbal communication, written text databases, and two speech databases. The trials to reveal the etymology of *izé* went back to the Finno-Ugric stems meaning 'thing' or 'something.' All the other explanations are admittedly vague, and there is no proper solution as to its origin. The present authors propose a novel idea that breaks away from traditional explanations. We assumed that the origin of *izé* was to be connected with the demonstrative pronouns (later definite articles) *ez/az* 'this' that occurred before a word the speaker wanted to produce but failed to find this word and needed some extra time for seeking for it. In addition, the articulation processes of the pronouns/articles (prolongation, shortening) and the factor of speech perception may lead to the formation of the present form of the sound sequence.

The paper discusses the history of this word from its first attested occurrence up to now. The first mention of *izé* was found in a dictionary in 1604. During the past centuries, *izé* occurred in various dictionaries (that provided various explanations for its use) and in other types of written texts referring to spontaneous speech (with diverse examples).

The word *izé* was analyzed in two speech databases: the Hegedűs Archive, which contains speech samples from the 1940s and 1950s, and the BEA spontaneous speech database, containing speech samples from 2009 and 2010. For the present study, the recordings of 36 speakers were used from the Hegedűs Archive. Nineteen occurrences of the target word produced by 16 speakers, both with and without suffixes, were found and analyzed (20 speakers did not produce *izé*). We found 162 occurrences of *izé* produced by 35 speakers (out of 50) in the spontaneous speech samples of the BEA database. For the use of *izé*, four categories were defined and analyzed: Word finding problem, seeking for the next thought, intended use of *izé*, and occurrences that could not be unequivocally categorized.

The use of the word *izé* has been stigmatized since relatively early on. The main reason for the strong stigmatization is that it calls the listener's attention to the speaker's speech planning problem. Despite this fact, we believe that *izé* will remain a part of the Hungarian speakers' mental lexicon for a long time as it is an extremely useful tool to help speakers gain time in case of speech planning trouble. This „filler” is unique and behaves like a linguistic chameleon in Hungarian.

Keywords: *izé*, etymologie, speech planning trouble, Hegedűs Archive, BEA database

1. Bevezetés

Az élő nyelv alapvető tulajdonsága a folyamatos változás a nyelv bármely szintjén és elemében. A nyelvi változás természetesen nem egyformán érinti a nyelvi szinteket, vannak könnyebben és lassabban, nehezebben és alig észrevehetően módosulók. A változásnak legnagyobb mértékben a szókincs van kitéve, amely gyorsan és jelentősen képes átalakulni (Wang et al., 2021; Gerstner, 2016). A nyelven belüli és a nyelven kívüli tényezők eredményeként létrejövő változások nem csupán új elemek megjelenését vagy eltűnését eredményezhetik, hanem a nyelvi elemek átalakulását is. Ez utóbbi többféle fokozatban jöhet létre. A szavak esetében a folyamat a fonetikai-fonológiai struktúra átalakulásához, olykor a morfológiai felépítés formálódásához, illetve a jelentés megváltozásához vezethet, ami gyakran a használati körnek és a stílusértéknek a módosulásával függ össze (Taylor, 2014). A módosulás és a változás rendszerint fokozatos, de mindenképpen ingadozásokat mutat a nyelvhasználatban.

A nyelvi változás mindig a nyelv használata során, a beszélő és a hallgató interakcióiban, jellemzően a beszédben jön létre. A változás alapja a beszélők állandó nyelvi alkalmazkodása a mindenkori kommunikációs és megismerési-megnevezési-kifejezési (kognitív) igényekhez (Kiss, 2016). Emiatt az egyes beszélők beszédében számos alkalmi újdonság jelenik meg akár fonetika-fonológiai, akár morfológiai, akár lexikai stb. szinten. Az újdonság elterjedését az teszi lehetővé, ha azt a beszélők egymáshoz alkalmazkodva egy kisebb, majd nagyobb közösség ismétlődő gyakorlatává teszik. Ekkor a változás következtében létrejött változat az adott kisebb vagy nagyobb beszélőközösség számára elfogadott norma lesz: a változás megtörténik. A magyar (Kiss, 2016) és a nemzetközi szakirodalom (Li & Siew, 2022) egyetért abban, hogy noha a változásokat a beszélők hozzák létre a saját nyelvhasználatuk során, azonban a változásnak többnyire nincsenek tudatában. (Kivételt képez például a nyelvi standardizáció szándékos normateremtése vagy a nyelvművelés, de a tipikus változás a kisebb nyelvi figyelemmel kísért nyelvhasználat eredménye.)

Említettük, hogy a változások, módosulások a nyelv bármely területét érinthetik, némelyek feltűnőbbek, mások kevésbé. Az e témakörben diakrón vagy szinkrón tekintetben folytatott vizsgálatok jó része a szókinccsel, a szavak jellemzőivel foglalkozik. Kutatásunkban egy olyan „szó” eredetével és használati sajátosságaival foglalkozunk, amely valójában nem is szó, tekintve, hogy a szó definíciója nem alkalmazható rá. Ez a „nem szó”, az *izé*, ami a nyelvhasználatban jól ismert, különleges, egyedülálló magyar lexikai elem. Mintegy hároméves kor

felett minden magyar beszélő mentális lexikonában megtalálható. Írott alakja az *izé*, kiejtve [ize:]¹, két szótagos, három fonémából áll, két elől képzett magánhangzóból és közöttük egy zöngés réshangból. Noha az *izé* megfelel nyelvünk hangrendi szabályainak, a hangsor mégis sajátos annyiban, hogy *-é*-re végződik². A mai magyarban ugyanis kevesebb mint 300 *-é* végződésű szó található (VégSz., 1941).

Miért érdekes az *izé* története és mai használata? Egyfelől azért, mert az elemzések talán választ adhatnak arra, hogy miért nem található hozzá hasonló az eddig vizsgált számos más nyelvben, másfelől mert kialakulásának a magyarázata a mai napig bizonytalan, és a használata önmagában is sajátos változásokat mutat. A nyelvi változások leírásában és magyarázatában általában ritkán kap szerepet a pszicholingvisztikai megközelítés. Tanulmányunk tárgya, az *izé* azonban nem nélkülözheti a hagyományos nyelvészeti elemzések mellett a pszicholingvisztikának, valamint a fonetikának a bevonását ennek az unikális magyar nyelvi elemnek a pontosabb megismeréséhez.

Az *izé* szónak (jobb híján „szónak” fogjuk nevezni) a furcsasága az, hogy nem valódi szó, hiszen nincs jelentése, nem hordoz semmiféle szemantikai tartalmat. A keleti nyelvjárásokban ugyan feljegyezték a népies orvosságnevek között egy *izé* szót ‘fehér ólompor’ (*pulvis cerussa*) jelentésben, de ez Benkóczy 1908-as gyűjtésén kívül sehol máshol nem szerepel, így aligha köthető ide.

Noha a „szó” fogalmának pontos definíciója máig vitatott a nyelvészetben (Haspelmath, 2023; Ladányi, 2017), abban egyetértés van, hogy a szó egyik legfontosabb tulajdonsága az, hogy rendelkezzen valamiféle jelentéssel. A töltelékszók a legtöbb nyelvben olyan valóságos szavak, amelyek alkalmilag az eredeti jelentésüket vesztő hangsorként használhatók, mint a magyarban a *szóval*, *hát*, *tudod*, *ugye* stb., avagy az angolban a *like*, *so*, *you know* (Laserna et al., 2014; Podlesskaya, 2010). Ezek a szavak az eredeti jelentésükben is használatosak (nem csak töltelékszókként), és diskurzusjelölői funkciót is kaphatnak, ezáltal pragmatikailag értelmezhető jelentéssel is rendelkezhetnek (Dér, 2020).

Mind a töltelékszók, mind a diskurzusjelölők számos tekintetben különböznek az *izé*-től. Ahogy említettük, az *izé*-nek nincsen saját lexikai vagy grammatikai jelentése, bár lehet pragmatikai funkciója. Tény, hogy a beszélők és a hallgatók számára is egyértelmű a szerepe a spontán megnyilatkozásokban. Az *izé* a magyar beszédben speciális töltelékszóként funkcionál, amely extra időt biztosít a beszélő számára a beszédtervezési nehézségek áthidalására. Természetesen nem kitöltött szünet, mint a hezitálás, és teljesen eltér a különböző diskurzusjelölőktől (Schirm, 2011). A beszéd során többféle lehetőség van arra, hogy extra időt nyerjünk a szavak, gondolatok, grammatikai szerkezetek megtalálására, avagy a beszédbeli

¹ A fonetikus átírást (IPA) csak a szükséges esetekben alkalmazzuk.

² A tanulmányban a beszédhangokat dőlt betűvel írjuk a könnyebb olvashatóság érdekében.

megakadások áthidalására, ilyenek például a néma vagy kitöltött szünetek, a szóismétlések, a nyújtások, de akár a torokköszörülés is.

Vajon miért lépnek fel beszédtervezési nehézségek, és miért van szükségünk különféle stratégiákra a megoldásukhoz? A spontán beszéd a beszélő gondolatait, megnyilatkozásának motivációit és céljait tükrözi, mondandójának megformálását grammatikai, fonológiai és fonetikai szinten, a hallgatónak címezve (Levelt, 1989). Beszéd közben nagyon sokféle cselekvést, mozgássort végzünk, beleértve a gondolatok kiválasztását, a nyelvi megformálásukat, a különböző pszicholingvisztikai, kommunikatív, szociológiai, érzelmi és egyéb döntéseket, valamint ezek ellenőrzését, míg a folyamat a kivitelezéshez, azaz a tágabb értelemben vett beszédprodukcióhoz vezet. A beszélő műveletei időben egyszerre zajlanak, ami azt jelenti, hogy egyidejűleg különféle beszédtervezési folyamatok történnek, és mindeközben a kiejtés is létrejön. Miközben beszélünk, a háttérben működnek a beszédtervezési folyamatok. Ezek a tények pedig mindenféle problémákhoz vezethetnek a beszédtervezés és a végrehajtás komplex jellegéből adódóan. Ha bárhol a mechanizmusban diszharmonia áll elő, a beszélőnek időre van szüksége a probléma felismeréséhez és megoldásához. Ezt az időszükségletet leggyakrabban a szünetek és a töltelékszók biztosítják.

Miért olyan különleges az *izé* a szemantikai tartalom hiányán túl? Nyelvünk agglutináló típusú, és igen gazdag morfológiával rendelkezik. Az *izé* többféle, különböző szófajú szóként jelenhet meg, lehet ige, főnév, melléknév stb., teljeskörűen felveheti az adott szófajra jellemző toldalékokat, és ezáltal a mondatban az adott szófajhoz tartozó összes szintaktikai funkcióba beilleszthető. Amikor a beszélő nehézséggel küzd a gondolat kifejezésében, a megfelelő szó kiválasztásában vagy előhívásában a mentális lexikonból, az *izé* segítségével könnyen meg tudja oldani a problémát. A megfelelő toldalék kiválasztása és alkalmazása révén a hallgató számára egyértelművé teszi, „kijelöli” a beillesztendő szó szófaját és szintaktikai szerepét a megnyilatkozásban. Mivel az *izé*-nek nincs önálló jelentése, a megfelelő morfológiai alakot felvéve, azaz toldalékolással bármiféle szemantikai tartalomra utalhat, annak hordozójává válhat. Kutatásunkban ennek a különleges elemnek a történetét követjük nyomon az első adatolásától, 1604-től több mint 400 éven át napjainkig, hiszen az *izé* ma is él a spontán magyar beszédben.

2. Módszerek

Vizsgálatunkban az *izé* előfordulásait, illetve a hozzá kapcsolódó megjegyzéseket kerestük a legelső adattól a szótárakban, nyelvtanokban, tudományos írásokban, szépirodalmi művekben (beleértve a színdarabokat és a novellákat), lejegyzett beszélt nyelvi anyagokban és két adatbázisban, amelyek spontán beszédet (is) tartalmaznak. Vizsgáltuk az *izé* előfordulásait és funkcióit a 20. század közepéről (Hegedűs Archivum), illetve a 2009-es, 2010-es évekből (BEA spontánbeszéd-adatbázis). Az *izé* egyes előfordulásának funkcióit minden esetben a jelen

tanulmány szerzői kategorizálták, kivéve, ha a hivatkozott mű szerzője ezt megtette.

3. Az *izé* etimológiája

3.1. Az etimológia problémái

Egy szó a nyelv belső változásának eredménye lehet, avagy más nyelv(ek) esetleges hatásaival magyarázható, amelyek a különféle kontaktusok révén jönnek létre a nyelv történetében éppúgy, mint a jelen állapotban. Az elemzés első kérdése, hogy miként illeszkedik az *izé* a magyar szavakra jellemző hangalakokhoz. A hangsor elöl képzett magánhangzókat tartalmaz, a szóvégi *é* azonban meglepő annyiban, hogy nem tekinthető jellemzőnek a magyar szavakra. Az *-é*-re végződő, relatíve kisszámú szavaink etimológiailag több csoportba oszthatók. Nagy részük idegen eredetű; idegen szavak vagy jövevényszavak, mint a németből (vagy legalábbis németből mint közvetlen átadó nyelvből) származó *karalábé* (< n. *kohlerabi*, *kalarabi*), *csokoládé* (< n. *Schokolade*), *parádé* (< n. *Parade*) stb.; a franciából *kommüniké* (< fr. *communiqué*), *rekamié* (< fr. *Récamier*) stb.; az ógörögből *szinekdokhé* (< óg. *szünekdokhé*), a latinból *roráté* (< l. *rorate*) stb. Egy másik részük valamilyen betűszó kiejtett alakjából származik: *tévé*, *ábécé*, *vécé* stb. Finnugor örökségünk az *-é* végű szavak csoportjában például a *lé* szó, amelynek *lev-* töve szabályos hangfejlődéssel lett *-é* végződésű (a *lev-* tö az *-s* képzős *leves* szóban maradt fenn). Az asszonynevek végén használatos *-né* elem szintén finnugor eredetű, a *nő* szavunk alakváltozata, valószínűleg a *neje* alak rövidülésével (*nej* → *nei* → *né*) jött létre. A belső keletkezésű határozószavak és névutói határozószavak, pl. *felé*, *mellé* *-é* végződése az ómagyar korban alakult ki a tövégi magánhangzóval összeolvadt *-i* lativusrag alkotta diftongusból. Van továbbá néhány ismeretlen eredetű *-é* végű szavunk, mint *purde*, *csalamádé* stb.

Az *izé* azonban egyik fenti csoportba sem sorolható biztonsággal, ezért eredetét igen sok találgatás övezte, és az etimológiája máig nincs megnyugtatóan tisztázva. Használata alapján nem valószínű, hogy az *izé* olyan átvétel lenne más nyelvből, mint az említett műveltségszavak, így belső keletkezésűnek vélhetjük. Az *izé* mai hangalakjának kialakulása az ómagyar korra valószínűsíthető. Korábbra azért nem, mert a szóvégi *-é* az ősmagyar kor vége előtt nem jelenik meg, viszont a középmagyar korban már biztosan számolhatunk az *izé* formával, lásd Szenczi 1604-es szótárát.

Az *izé* eredetét már az elődeink is igyekeztek megmagyarázni. Az első kísérlet a Czuczor–Fogarasi-szótárban található. Czuczorék az *izé* eredetével kapcsolatban azt feltételezték, hogy az *ez* mutató névmással van kapcsolatban, „rokon vele *ez*, s ennek mintegy kettőztetéséből származottnak látszik: *ez-e*” (CzF. 3: 161). További magyarázattal a szótár nem szolgált, sem azzal kapcsolatban, hogy miért történt a feltételezett megkettőződés, illetve hogy miként alakult volna ki ebből a hangsorból az *izé*.

Mivel a magyar finnugor nyelv, értelemszerűen a rokon nyelvekben is próbáltak hasonlóságokat találni. Az egyik elképzelés szerint az *izé* a finnugor *ut*, *ot*, *atta* tövekre vezethető vissza (Szinyei, 1911; SzófSz., 1941), amelyek jelentése ‘dolog’, illetve ‘valami’, és a szó végén álló elemet valamiféle képzőnek vagy szónyújtó toldaléknak tartották. A feltételezés szerint az *izé* valaha valódi szó volt ‘valami’ jelentéssel (és egyéb, hasonló jelentésű névmások szemantikai tartalmával). A szó alakjának magyarázata azonban nem volt kielégítő, mint ahogyan az sem, hogy miért éppen az *izé* veszítette el a feltételezett jelentését más névmásokkal szemben.

Prohászka (1957) elfogadta ezt az etimológiát, és igyekezett megmagyarázni a szóvégi *-é* magánhangzót is. Úgy gondolta, hogy az eredeti alak *izei* vagy más nyelvjárási változatban *izej* volt, a szó végén egy kicsinyítő képzővel, szerinte az utolsó két szegmentum egy hosszú *é* magánhangzóra olvadt össze.

Egy másik magyarázat az *izé* és a ‘betegség’, ‘rák’ jelentésű *íz* szó között próbált kapcsolatot kimutatni. Simonyi (1876) a hetési és somogyi *izej* alakot tartotta az eredetibbnek, és az ‘ördög, fene’ jelentésű *iz* szót látta benne, meg egy ehhez járuló ugor alapnyelvi *-j* deminutív képzőt. Többféle alakot is idevont (*eperj*, *kicsi* stb.), ezek azonban mai tudásunk szerint részben tévesek (a *kicsi* az ótörök *kicsiny*-ből rövidült), részben nem kicsinyítő képzős formák (az *eperj* szóvégi *-j*-je egy fgr **-k* denominális névszóképzőből ered). Ami az *-i* kicsinyítő képzőt illeti, az ESz. szerint „az *-i* kicsinyítő képző valószínűleg több formáns (ti. forma) egybeesésének köszönhető. A nyelvész és fonetikus Balassa 1940-ben szintén az *izé* és a ‘betegség’, ‘rák’ jelentésű *íz* szó összefüggéséről írt. Magyarázatot azonban nem adott hozzá.

Egy harmadik megközelítés az etimológia tekintetében Juhász Jenő 1959-es cikke. Azt írja, hogy „Ez az *iz* tehát ‘ez’ jelentésű, az *i* pedig emphatikus nyúlás eredménye, hiszen ingerült állapotban gyakran megnyújtjuk a károm szó [káromkodás] első tagját: »Az örrdög vigye el!« »A ffene egye meg!« »Asszta mindenét neki!« Ugyanígy: »Az íz egyen meg!«”

A modern etimológiai szakirodalom – TESz. (6), EtSz. (7: 633), ÚESzWeb – szerint az *izé* eredete bizonytalan, ismeretlen, avagy finnugor eredetre mutat. A TESz. szerint összetett vagy származékszó, a tő rokon nyelvi párhuzamai *at-*, *ot-*, *ut* alakúak, az *-é* végződés pedig a mutató névmási eredetű *-e/-é* kérdőszóra vagy egy *-é* (< *-ei*, *-ej*) képzőre vezethető vissza, azonban az EWUng. és az ÚESzWeb rámutat az eredetileg hátul képzett kezdő magánhangzó (fgr **a > i > i*) és az *-é* végződés közötti ellentmondásra.

Bár a szótárak a jelentésfejlődést expliciten nem tartalmazzák, az általános szóhelyettesítő jelentésű szavakkal való rokonítás, az azokból való kiindulás azt feltételezi, hogy az etimológiai szótárak logikája szerint a ma használatos kétféle funkció közül a szópótló volt az elsődleges, és abból fejlődött ki a beszédprodukció valamiféle megakadására utaló használat.

Ugyanakkor az *izé*-re vonatkozó feljegyzések kezdettől egyértelműen a megakadást, a szótalálási nehézség áthidalását emelik ki. Az a megállapítás, hogy

az *izé* egyfajta töltelékelem (*transcendens*), ismereteink szerint először Geleji Katona (1645) megfogalmazásában található (lásd munkája 34. pontjában, alant részletezve). Ugyanígy a szótalálási nehézséget emeli ki Gyarmathi (1794, 1799), a Kresznerics-szótár (1831), a Czuczor–Fogarasi-szótár (1863) és Ballagi szótára (1872).

Láttuk, hogy az ismertetett etimológiai magyarázatok egyike sem alkalmas az *izé* etimológiájának egyértelmű felderítésére. Ezért úgy véljük, érdemes új szemszögből vizsgálni a kérdést. Mi történik akkor, ha onnan indulunk, amire a legkorábbi feljegyzések is utalnak, vagyis a megakadás felől? Az áttekintett anyag alapján, a meglévő eredetpróbálkozások, valamint a spontán beszéd folyamatának ismeretében felvázoljuk az *izé* kialakulásának egy lehetséges útját. Noha az elméletünk spekulációnak tűnhet, a folyamatot a kiejtés, a beszédészlelés, a valószínűsített elterjedés tényeinek és a nyelv, illetve a nyelvhasználat működésének figyelembevételével hoztuk létre.

3.2. Egy új hipotézis az *izé* kialakulására

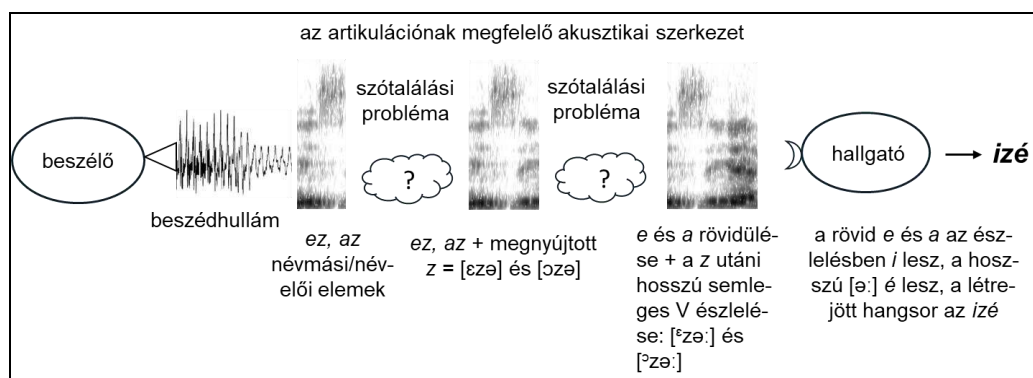
Mi van akkor, ha az *izé* valós eredete valójában nem egy szó kialakulásához kötődik, vagyis elképzelhető-e, hogy soha nem is volt semmilyen jelentése (lexikális szemantikai tartalom értelmében, amint az a valós szavakra jellemző)? Tételezzük fel, hogy a sok-sok évszázaddal korábban élt beszélők ugyanúgy szembesülhettek a spontán közlések létrehozásának számos problémájával, mint a mai beszélők. Valószínűsítsük továbbá azt, hogy kialakult egy hangsor, amellyel bármilyen szót tudtak helyettesíteni a közléseikben, sőt – talán később – arra is alkalmassá vált, hogy általános töltelékszóként többletidőt biztosítson a beszélőnek a beszédtervezési folyamatban.

A hipotézisünk a beszélő jellemző viselkedésén alapszik, ami akkor fordul elő, amikor a spontán közlések során szótalálási nehézségbe ütközik. Tételezzük fel, hogy a beszélő kiejti a mutató névmási (később: határozott névelői szerepű) *az*-t vagy az *ez*-t a kiejtendő szó előtt, azonban hirtelen nem találja a kívánt szót a mentális lexikonában. (Az *az/ez* határozott névelő a mutató névmásból alakult ki a magyar nyelvtörténetben, az *a* alakváltozat kialakulása későbbi.) Mi történik ebben a beszélő számára kellemetlen helyzetben, hogy ti. nem jut eszébe a kiejteni szándékozott szó? A beszélő a mentális lexikonból előhívni kívánt elemre a mutató névmás (később a belőle alakult határozott névelő) használatával mintegy rámutat, miközben keresi a szót (egyúttal aktivizálva a hallgató mentális lexikonját is, akiről feltételezi, hogy szintén ismeri a közlésébe szánt, de még nem kiejtett szót). Eközben a beszélő mintegy akaratlanul megnöveli a névmás/névelő második beszédhangját, a *z* mássalhangzót. A nyújtás mint időtöbbletet biztosító stratégia jól ismert ilyen helyzetekben, napjaink spontán közléseiben is (Gósy & Eklund, 2018). A megnyújtott mássalhangzó időtartama lehetőséget ad a beszélőnek az ejteni kívánt szó további keresésére a mentális lexikonban. A zöngés *z* réshang nyújtásának artikulációs következménye, hogy létrejön egy neutrális, semleges magánhangzó, a *svá* ([ə]), ami egy *ö* magánhangzóra

emlékeztető realizáció. Az eredeti névmások/névelők ennek megfelelően az észlelésben már három-három szegmensből állnak, az [ɔzə], illetve az [ɛzə] hangsorokból.

Mi történhetett a továbbiakban? A beszélők elkezdtek ösztönösen használni a nyújtott mutató névmásokat/névelőket tartalmazó hangsort időnyerési célból, valahányszor nem találták meg a kiejteni kívánt szót a szükséges pillanatban. Az ismételt ejtések odavezettek, hogy rövidült a hangsorkezdő magánhangzó, és meghosszabbodott a zöngés réshang nyúlásával keletkezett utolsó magánhangzó, a semleges [ə]. A kétféle időtartam-változás eredményeképpen létrejött hangsorok így hangozhattak: [ɛzə:], illetve [ɔzə:]. Végezetül a beszédészlelés folyamatában rögzült a hangsor végleges formája. A megrövidült első magánhangzó az *i* benyomását keltette (vö. Gósy, 1989), a mássalhangzót követő hosszú semleges magánhangzó pedig az *é* hangra emlékeztetett, amiben a réses ajakartikulációnak is nagy szerepe lehetett (1. ábra). A mutató névmások/névelők egy adott beszédtervezési helyzetben, a szótalálási probléma hatására átalakultak, fonetikai formájuk módosult, és létrejött a szóra emlékeztető *izé* hangsor.

1. ábra. Az *ez*, az névmási/névelői elem átalakulása az *izé* hangsorrá szótalálási probléma hatására



Az *izé* kialakulásának egyes „fokozatai” egyértelműen alátámaszthatók artikulációs, akusztikai-fonetikai, beszédészlelési és pszicholingvisztikai paraméterekkel. Noha az *izé* jelentés nélküli, funkciója azonban van, és valószínűsíthetjük, hogy meglehetősen gyorsan kezdett terjedni a beszélő közösségekben. Az, hogy a hangsorhoz nem tartozott szemantikai tartalom, nem jelentett gondot a kommunikációban, hiszen a funkciója egyértelmű volt a nyelvhasználók számára. A beszédprodukciós és beszédészlelési hurok, azaz a két folyamat összefüggése, illetve egymásba fonódása tette lehetővé, hogy valódi szavakhoz hasonlóan, a nyelvhasználók tárolják az *izé* akusztikai-fonetikai reprezentációját a mentális lexikonukban a hozzárendelődött funkcióval együtt.

Az *izé* elterjedése a következő tényezők mentén magyarázható: (i) a töltelékszói funkció, amely kiváló stratégia a beszédtervezési nehézség feloldására, (ii) egyszerűen kivitelezhető artikuláció, (iii) bármely szófajt helyettesíthet és (iv) bármilyen toldalékkal ellátható, ha arra szükség van. Ez

utóbbi két tényező elősegítette a hallgató számára annak valószínűsítését, hogy milyen szót kereshetett a beszélő, ami könnyítette a beszédmegértést. A funkció, vagyis az időnyerési stratégia alkalmazása az *izé* kiejtésével nyilvánvalóan hamar kiszélesítette a használati körét. Az egyszerű szóhelyettesítés mellett egyéb beszédtervezési nehézség, diszharmónia esetében is kezdték használni, amikor a beszélőnek többletidőre volt szüksége, és folytatni kívánta a közlését. A stratégia elterjedése és a használat bővülése egymást erősítették.

A fentebb részletezett hipotézisünk találgatásnak (is) tűnhet, azonban meg vagyunk arról győződve, hogy a jelen ismereteink és tudásunk szerint ez a feltételezés látszik a legvalószínűbbnek az *izé* eredetére és kialakulására vonatkozóan. A hipotézisünk realitását további, megkérdőjelezhetetlen tények is erősítik. (i) A hangsor mindhárom beszédhangja része a magyar nyelv fonémaállományának legkésőbb az ómagyar kor vége óta. (ii) Az *izé* első említéskor a magyar nyelv történetében két mutató névmás/határozott névelő volt, amelyek csak a kezdő magánhangzójuk minőségében különböztek, és amelyek kiejtéséhez köthető az *izé* létrejötte. (iii) Mindkét magánhangzó redukálódása ugyanolyan azonosítást eredményez a beszédészlelésben, esetünkben az *i* magánhangzót (Gósy, 1989). A redukálódás könnyen magyarázható azzal, hogy a névelői szerepű névmáson nincs hangsúly, ami kedvez a rövidülésnek. (iv) A kutatási adatok több nyelv esetében is igazolták, hogy a beszélő például a névelőt akkor is kiejti, ha még csak a lemma szintre jutott el a mentális lexikon aktiválásában (Jescheniak & Levelt, 1994). (v) A zöngés réshang (z) nyújtása voltaképpen akaratlan, az idézi elő, hogy a beszélő ezzel is igyekszik magánál tartani a szót, mintegy jelezve a közlés folytatását, miközben igyekszik megtalálni a kiejteni kívánt szót. A spontán beszédben tapasztalható nyújtások stratégiája ugyanígy működik. A beszélő a megelőző kiejtett szót, esetünkben a névmást, később névelőt, megnyújtja, ami számos más, vizsgált nyelvhez hasonlóan, a magyarban is jellemzően a hangsor végén történik, az előfordulások több mint 60%-ában (Gósy & Eklund, 2018). (vi) A társalgás során a hallgató könnyen felismeri az elhangzó három beszédhangból álló hangsor funkcióját, és beszélőként átveszi a stratégiát, vagyis alkalmazza az *izé* hangsort beszédtervezési nehézség esetén.

Megjegyezzük, hogy különböző névmásokat és a 'dolog' jelentésű szavakat gyakran használják a beszélők időnyerő funkcióban más nyelvekben, de azok eredeti jelentése mindig megőrződött (Podlesskaya, 2010). Podlesskaya töltelékszóként felsorolja a magyar *izé*-t is, azt azonban nem részletezi, hogy ennek a hangsornak valójában nincs lexikai-szemantikai jelentése szemben az összes többivel.

Az elmondottak alapján hipotézisünket az *izé* létrejöttére vonatkozóan bizonyítottnak tekintjük, belső nyelvi fejleménynek tartjuk, és az ómagyar korra datálhatjuk.

4. Az *izé* írásbeli előfordulásai

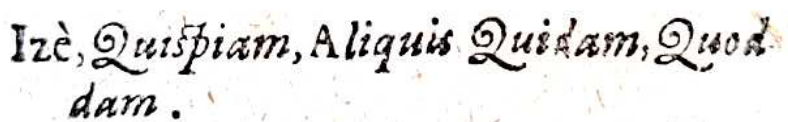
4.1. Az *izé* a szótárakban és a nyelvészeti munkákban

A gyakran használt szavak hangalakja hosszú időn keresztül megőrződik (Pagel et al., 2007). Joggal feltételezhetjük, hogy az *izé* kezdettől a gyakran használt szavak közé tartozott, fonetikai formája nem változott, és a beszélői attitűd, vagyis az időnyerési szükséglet kielégítése révén gyakorlatilag azonnal beépült a hangsorba (ez egyfajta szubjektifikáció, vö. Hamilton et al., 2016).

A szótárak általában az adott beszélőközösség által használt szavakat tartalmazzák. Egy szó megjelenése a szótárban feltételezi, hogy azt széles körben ismerték és használták. Korai szótáraink a magyar szókincset a latinhoz (illetve más nyelvekhez) igazítva rendszerezték. E szótárak ezért nem csak az aktuálisan használt mindennapi szókincs dokumentálását jelentették. Szerkesztőik az iskolai oktatást szem előtt tartva az (antik és egyházi) műveltségi kánon szókincsét is belevették a szótárba, sőt, aktívan törekedtek a nyelvi eszmény kialakítására, a polifunkcionális és intellektualizált irodalmi nyelv megteremtésére (Szathmári, 2007). Ebből kiindulva a jellemzően a szóbeliséghez kötött, beszédtervezési, szótalálási, szóhelyettesítő problémákhoz kapcsolódó elem felvétele a szótárakba arra utal, hogy az *izé* gyakran és hangsúlyosan volt jelen a beszélők kommunikációjában. A szótár a műfaj természetéből adódóan a latin vagy német vagy egyéb nyelven megadott fordítással csak a szóhelyettesítő funkciót képes visszaadni, amikor az *izét* általános vagy határozatlan névmással fordítja. Ugyanakkor a nyelvművelő irodalom (Geleji Katona, 1645) már igen korán regisztrálja a beszédtervezési zavart áthidaló időnyerő funkciókat is, és a későbbi szótárakban is megtaláljuk a töltelékszó funkció leírását.

Az *izé* első említését Szenczi Molnár Albert *Dictionarium Latino-Ungaricum et Ungarico-Latinum* című, 1604-ben megjelent szótárában találjuk. A 2. ábrán látható a szóban forgó címszó: *Izé. quispiam, aliquis, quidam, quoddam* ‘valaki, bárki, egy bizonyos valaki, valami’. (Megjegyzendő, hogy a mai magyar–latin szótárakban az *izé* fordítása már nem névmás, hanem *res* ‘dolog’.)

2. ábra. Az *izé* első említése Szenczi Molnár Albert 1604-es *Dictionarium*ában



Izé, Quispiam, Aliquis, Quidam, Quoddam

Mintegy 150 évvel később, a Szenczi-szótáron alapuló Pápai Páriz-szótárban, amit Bod Péter némettel egészített ki (Pápai Páriz & Bod: Dict.), az *izé*-re már német értelmezést is találunk: *jemand, etwan, einer*. 1831-ben Kresznerics Ferenc szótára (1831) a 266. oldalon Szenczit ismétli meg. A szerző megemlíti az *izé* szótalálási nehézséget áthidaló funkcióját: „Pótoló szó, midőn valamelly dolognak a' neve nem jut hirtelen eszünkben, és ekkor IZÉvel szoktunk

magunkon segíteni.” Kresznericsnél megjelenik címszóként a képzett ige is: *izél*, *izéli*, példamondattal: *Mit izélsz annyit, Jól meg izélték 's a' t.*

A magyar–német szótárak (Kresznerics, 1831; Anon. szerző, 1838; Fogarasi, 1860) és Loos József magyar–német–tót (ma: szlovák) szótára (1869–1871) az *izé*, *izél* szó német megfelelőit így adják meg: *das Ding, irgend, etwas, ein gewisser, machen, tun*. Az egynyelvű Czuczor–Fogarasi-szótár noha ad értelmezést is ('valaki, valami'), alapvetően a spontán beszéd során tapasztalható szótalálási nehézségeket emeli ki, és ezeket stigmatizálja. A feltehetően a szerzők alkotta, itt-ott mesterkelt példamondatok azt kívánják bemutatni, hogy az *izé* többféle szófajként is viselkedhet: „Pótoló szó, s határozatlan értelmű, melylyel élünk, midőn valamely dolog neve hirtelen nem jut eszünkbe, s am. *valaki, valami. Hát az izét elhoztad-e? Tegnap izével találkoztam. Mondd meg izének, hogy. Alakúl belőle melléknév is: izés. Ezen ember egy időtől fogva olyan izés. Továbbá határozó: izésen. Olyan izésen áll a kalapja. Végre ige: izél. Ugyan ne izélj már nó. Mit izélsz? Izéné megizélte az izéjék gyerekét.* Ily beszédmód a tunya gondolkodás jele, s mind írásban, mind mivelt beszédben kerülendő. Helye legfőlebb tréfás népjellemzésben lehet” (CzF. 3:162).

Ballagi az 1872-es *A magyar nyelv teljes szótárában* mind főnévként (*izé*), mind igeként (*izél*) feltünteti a szót. Ő is szótalálási nehézségekhez köti: „az a bizonyos dolog, v. személy, melyről, v. illetőleg kiről szó van, de melynek, v. kinek neve nem jut eszünkbe” (1872: 635). Példáiban főnévi, igei, melléknévi, igenévi, határozószói szerepben is szerepel: *nem láttad izét, a bátyádat?; mondd meg izének, hogy jőjön el délután* stb. A szótárban először találjuk meg az *izél* 'növel nemileg közösül' értelmezését, itt egyértelműen a trágár szó helyettesítéséről van szó. Érdekesség, hogy mind Ballaginál (1872: 85), mind a Czuczor–Fogarasi-szótárban (CzF. 1: 520) megtalálható az *izél*, mint az erdélyi *békálkodik* címszó értelmezése 'oktalanul kötekedik', 'izél'.

A szótárak mellett értékes, bár mozaikos forrást jelentenek a grammatikák és a nyelvművelő munkák is. A legkorábbi említés egy 17. századi nyelvművelő-nyelvhelyességi munkában, Geleji Katona István 1645-ös *Magyar Grammatikatska* című művében található. Szerinte az *izé* töltelékszó, amely beszéd során előálló megakadást, pontosabban szótalálási nehézséget jelez. A nyelvművelő írás az *izé* használatát a nyelvi hibák között említi. „XXXIV. „*Izé* mint egy transcendens a mi nyelvünkben: mert akármit, ha hamar eszében nem jut, a magyar mind *izének* mond: *Amaz, izé, így szollván; vagy. izét; s' a' tselekedetet megént izélfnek: Mit izélsz?* ugymond. Olyan a' dolog-is,

3. ábra. Az *izé* töltelékszó funkciójának első említése Geleji Katona Istvánnál 1645-ben

XXXIV. *Izé* mint egy transcendens a mi nyelvünkben; mert akármit, ha hamar eszében nem jut, a Magyar mind *izének* mond: *Amaz, izé, így szollván; vagy. izét; s' a' tselekedetet megént izélfnek: Mit izélsz?* ugymond. Olyan a' dolog-is,

Itt érdemes megjegyezni azt is, hogy az idézet végén van egy lényeges mondat. Geleji a *dolog* szót adja meg az *izé* szinonimájaként: *Ollyan a' dolog-is*. Ez az értelmezés azonban csak később jelenik meg a szótárakban (das Ding, mai latinszótárak: res). Szenczinél és a Pápai Páriz-szótárban csak a határozatlan névmások találhatók értelmezésként latinul és németül is, mint *jemand, aliquis quoddam, etwan* stb.

150 évvel később, Gyarmathi Sámuelnél találjuk meg az *izé*-t az *Okoskodva tanító Magyar Nyelvemester* című 1794-es grammatikájában. Gyarmathi világosan leírja az *izé* többféle szerepét is: időnyerést a szótalálási nehézség feloldására, illetve a szóhelyettesítést, amikor a beszélő nem tud vagy nem akar valamit pontosan kifejezni: „Ezzel a' szóval a' beszéd minden részei helyett lehet élni ... Mikor ezt a' szót *izé* mondjuk, még a'kor az a' dolog a' mellyről akarunk szólni, vagy egészen nem jutott eszünkbe, hanem tsak a képzelődésünkben vagy, vagy pedig nem tudjuk vagy éppen nem-is akarjuk világosan ki-fejezni” (Gyarmathi, 1794: 211). Számos példát, rövid társalgásrészletet is hoz: *Ne izélj!; Láttam az izéjét.; Apád izékor jött haza*. Hangsúlyozza továbbá, hogy az *izé* bármilyen szófajt helyettesíthet, ami semmilyen más nyelvben nem található jelenség. Ugyanez a szerző a magyar nyelvet más finnugor nyelvekkel összevető munkájában, az 1799-es *Affinitas*ban megismétli ezeket az állításait, és a lapp [ma: számi] *ata* névmással állítja párhuzamba az *izét*. Gyarmathi így fogalmaz az *Affinitas* 51. oldalán: „lapp *ata*. at. magyar *Izé* lapp *Atet*. magyar *Izélni*. ... ehhez hasonló egy európai nyelvben sincs. Amit a németek az ő *das Ding*-ükkel igyekeznek kifejezni, az a látszat szerint ezt valamennyire megközelíti. De a lappok *ata*-ja és a magyarok *izéje* a melléknév, a főnév, az ige, a melléknévi igenév, a határozószó feladatkörében is szerepet vállal. Lássunk példákat! *Hallod-é, Izé! Gyere bé!* Itt a beszélő Pétert, Jánost stb. akarja megnevezni, de az illető neve nem jut eszébe stb.” (1799: 51).

A következő megjegyzés a 19. század közepéről való, Brassai Sámuel 1864-es nyelvészeti munkájából, és számunkra azért érdekes, mert világosan leszögezi, hogy az *izé* használata az élőbeszédekre korlátozódik: „...van a magyar élő beszédben egy szó, mely minden névnek képviselő synonymája: 'izélés' ennek egy származéka: 'izél', mely hasonlóképpen minden kitelhető ige helyett szolgál” (1864: 211).

Vele egy időben Budenz ismét idegen nyelvi párhuzamokat keresett. Egyik írásában a csuvas *hajghi* névmás magyarázatával foglalkozik. Ezt a szót Reguly Antal korábban úgy fordította, hogy „azonos (sic) azon dolog, melyről beszéltünk, de melyet nem akarok nevezni”. Budenz (1862: 268) szerint e csuvas névmás használata a magyar *izé* használatához hasonlatos: „beszéld el, mit *izékor* láttunk (hiszen tudod mikor), meg: addsa ide az *izét* (hiszen tudod mit)”.

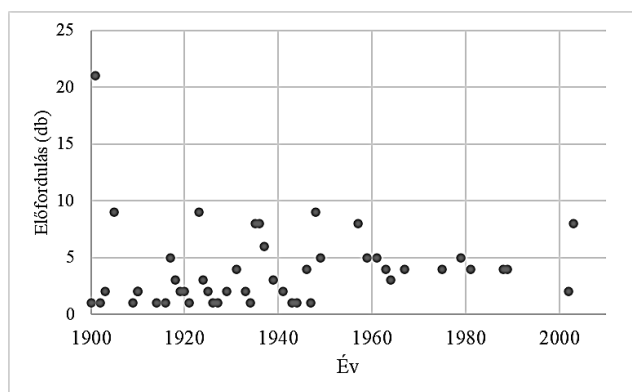
4.2. Az *izé* az írott szövegekben

Noha használatát, mint láttuk, már a 17. század elején regisztrálták, az *izé* egészen a 18. század végéig feltűnően hiányzik a ránk maradt szövegekből. Hiányát az magyarázhatja, hogy a fentebb felsorolt szótárak, grammatikák, nyelvművelő munkák tanúsága szerint mindig is a szóbeliséghez, ezen belül a spontán beszédhez kötődött, ráadásul kezdettől stigmatizált volt. Véleményünk szerint ez a korai stigmatizáltság voltaképpen a beszélőt érintette. Nem az *izé* hangsort kifogásolták tehát, hanem azt a helyzetet, hogy a beszélő nem találta meg a megfelelő szót, amit ki akart ejteni, vagy nem tudta megakadás nélkül folytatni a gondolatot. A beszédtervezési diszharmóniát a beszélő hibájául rótták fel. Az *izé* stigmatizáltsága tehát nem a hangsorból vagy a funkcióból következett, hanem magát a diszharmónia jelenségét tekintették egyfajta retorikai „hibának”.

A beszélt nyelvi szövegek írott lejegyzése során az írnokok, deákok nyilván nem rögzítették a megakadás feloldására használt töltelékszót, a levelek, egyéb szövegek fogalmazása közben pedig nem volt szükség rá, lévén elég idő állt rendelkezésre a szövegtervezésre, szótalálásra. Az *izé* felbukkanása ezért az élő beszédet szándékosan utánzó szövegekben várható.

Az ómagyar korból igen kevés, elsősorban vallásos tárgyú és műfajú szövegemlékünk maradt fenn az Ómagyar korpuszban, és ezekben – a fentiek miatt – értelemszerűen nem várható egy ilyen elem megjelenése. Az élő nyelvhasználatához és a szóbeliséghez közelebb álló, ómagyar és főként középmagyar kori szövegeket tartalmazó Történeti Magánéleti Korpuszban sem találjuk meg, holott a boszorkányperek tanúvallomásaiban elképzelhető lett volna a megjelenése. Hasonlóképpen hiányzik a Középmagyar emlékirat- és drámakorpuszból, holott a drámák szövegében szintén valószínűsíthető lett volna. Előfordul viszont Czeglédi István református lelkész 1669-es, Sámbar Mátyás jezsuitával hitvitázó könyvében (1669: 219), ahol élő nyelvhasználatot, elképzelt párbeszédet imitál a szerző. Ellenfele érveit gúnyosan minősítve ezt mondja. „Tudva kel már Sámbart bölcs tudományáért (*izé!*) tartani.” Az *izé* itt egy trágár szót hivatott pótolni.

Az *izé* az újmagyar kortól jelenik meg nagyobb számban a szövegekben, főként a népszínművekben és népies elbeszélésekben, amelyekben – ahogy Czuczorék rámutattak – a „tréfás népjellemzés” eszközeként használták, vagyis a nehezen fogalmazó, iskolázatlan szereplők spontán beszédének utánzására. Az *izé* újmagyar és újabb magyar kori előfordulásait a Magyar Történeti Szövegtár 1772 és 2000 közötti, különféle műfajú és terjedelmű nyomtatott szövegein vizsgáltuk, amelyek reprezentatív áttekintést adnak a kor irodalmi, köznyelvi és tudományos nyelvhasználatáról. Az 1770 és 1950 közötti időszakot vizsgálva az *izé* 248-szor fordul elő, először 1786-ból adatolható. A gyakorisági eloszlása változó, évente 1-2 alkalommal található meg (4. ábra). (Az adatokat torzíthatja a szerzők egyéni szóhasználat is, például ugyanazon szerző 1889 és 1901 közötti szövegeiben 21 alkalommal használta az *izé* szót.)

4. ábra. Az *izé* előfordulása az 1900-as évektől 2003-ig (a Magyar Történeti Szövegtár alapján)

Példák a Magyar Történeti Szövegtárból: *aki itt izén is lakott sok ideig, hun csak no, ejnye no, bizon nem jut eszembe, ej pedig nyelvem hegyén van, a patvar üsse meg* (1836); *Izé, mit is akartam mondani?* (1889–1901); *A szegény leány házát – követem alásan – izével egészen környös körül hordta*. Az elkerült szó feltehetőleg a *gané* vagy a *szar* szó lehetett.

A magyar nyelvjárástanhoz közlő szöveget Vass József 1864-ben. Itt fordul elő az *izé* a következő kontextusban: „... Ki is mondá a minap rólad? – *izé* fogá mondani – no, kivillant a fejemből, hogy te ojan kényeske vagy...”. Értelmezésünk szerint a tulajdonnév nem jut a beszélő eszébe, ezért használja az *izé* szót, erre utal a folytatásban, hogy „kivillant a fejemből”, vélhetőleg a név (1864: 18).

A különböző műfajú fordításokban is megjelenik az *izé*, az eredeti magyar művekhez hasonlóan a szereplők vagy a helyzet jellemzésének céljával. Egy török népszínmű fordításában Kunos Ignác (1887) azért használja az *izé* szót, hogy érzékeltesse a szereplő zavarát, aki túl bizalmasnak érzi, amit kérdez: *kérdeznéd tán tőlük hogy izé ura van-e*. Kétszer fordul elő az *izé* egy 75 esztendőös nyugat-magyarországi parasztember 1903-ban elhangzott elbeszélésének lejegyzésében (Varga, 1908). Az egyik ebben a mondatban: *Od vótam eh hétig ebbe a izébe, onnan lékisérteg bennünket Világosra, mast osztán gyalogok marattunk, lú neküm marattunk*. Ebben a közlésben a beszélő nem tudja felidézni a hely nevét, ahol volt. A másik előfordulás is szótlalálási problémára utal: *hát mink ott töprenkettünk a magunk izejibü* (talán az *eszéből* szót helyettesíthette). Egy 1983-as nyelvjárási szövegrészletet közölnek a *Magyar Nyelvőr* lapjain (K. Fábián, 1986: 316), amely több *izé*-t is tartalmaz, például *Hát osztá akkor ez izéba vöt, negyvenháromba* (1986: 318). Az irodalomban további példa Rejtő Jenő egyik regénye (*A Láthatatlan Légió*), amiben az *izé* használatával jellemzi a főhős verbális esetlenségét és furcsa személyiségét (ezt a regényt kizárólag azért emeltük ki, mert az *izé* rendkívül gyakran hangzik el a főhős beszédében, és reprezentálja az *izé* ejtésének többféle okát).

Az írott nyelv adataira érdemes kitekinteni. A Hungarian Web 2023 (huTenTen23, letöltve 2023 novemberében) lekérdezésének eredménye szerint –

4 320 782 243 szó alapján – az *izé* összesen 41 440 előfordulást mutat, ami azt jelenti, hogy egymillió szóban közel 10 db található. A toldalékoltak aránya az összes *izé*-előfordulás tekintetében mintegy 4 db. Noha a teljes számennyiséghez képest ez elenyészőnek tűnhet, a negyvenezerenél is több előfordulás mégis egyértelműen aktív használatra utal.

Elemzésünk következő szakaszában két spontánbeszéd-adatbázisban kerestük az *izé* előfordulásait, és vizsgáltuk a funkcióit.

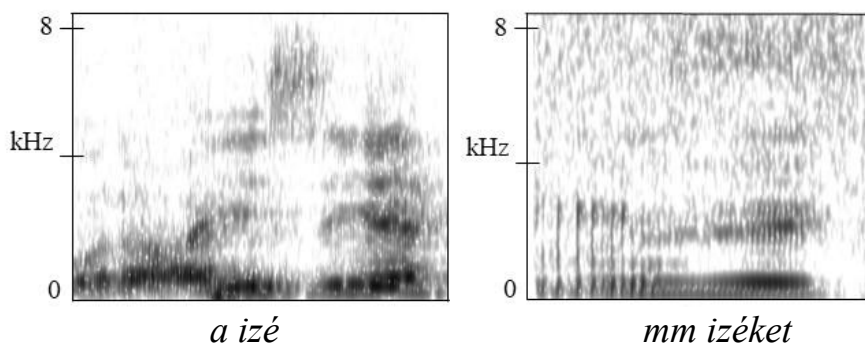
5. Az *izé* szóbeli előfordulásai

5.1. Az *izé* a Hegedűs Archívumban

Az 1940-es, 1950-es évekből származik az első olyan beszédatadtbázis, amely nagy mennyiségű, különböző műfajú hangzó magyar szöveget tartalmaz (Gósy et al., 2011). A kiváló nyelvész és fonetikus, Hegedűs Lajos nevéhez fűződik a nyelvjárási beszéd rögzítésének megkezdése mintegy 85 évvel ezelőtt. Célja az volt, hogy különféle stílusú beszédanyagokat (spontán szövegek, például lakodalom, disznóölés elmesélése, mesék, betyárhistóriák, katonatörténetek, verbuválás, tréfák, gyermekcsíny, búcsú stb.) rögzítsen az ország különböző pontjain. A gyűjtések Magyarország területének számos megyéjében, valamint határon túli helyeken történtek. A teljes anyag több mint 1700 felvételtől áll. Összesen 50 órát meghaladó hanganyagot rögzítettek, a 600 adatközlő többsége idősebb volt 40 évesnél.

Ez az adatbázis lehetőséget ad az *izé* előfordulásainak vizsgálatára. Összesen 36 beszélő (20 nő és 16 férfi) beszédanyagát (mintegy 4,3 óra) elemeztük nyolc különböző felvételi pontról. Az 5. ábra két hangszíneképet mutat, amelyek az *izé* ejtéséről készültek (a toldalékolt *izé* spektrogramján látható zajnyalábok a felvételi körülmények következményei). A két – hangszíneképekkel is illusztrált – példában a szótalálási nehézséget az *izé* használatán túl a névelő nyújtása, illetve a megelőző *mm* kitöltött szünet is előjelzi.

5. ábra. Az *izé* hangszíneképei két női beszélő ejtésében (a vízszintes tengely az időt mutatja). A bal oldali képen az *izét* nyújtott *a* névelő, a jobb oldali képen az *izéket* hangsort pedig egy nazális mormogás előzi meg



A 36 beszélő közül 16-an produkálták a keresett hangsort toldalékos vagy anélküli formában (20 beszélőnél egyáltalán nem fordult elő). A narratíváik összesen 19 előfordulást tartalmaztak, adatközlőnként 1-2 ejtést, óránként mintegy 4,4 *izé* hangzott el. A közvetlen funkciót tekintve 14 előfordulás (73,7%) szótalálási nehézségre utalt, 5 pedig (26,3%) gondolati szintű beszédtervezési diszharmónia esetén biztosított többletidőt a beszélőnek. A szófajok tekintetében, az *izé* 10 esetben főnevet (71,4%), 3 esetben (21,5%) igét és egy esetben (7,1%) számnevet „helyettesített”. A szótalálási nehézség kialakulásakor egy kivétellel valamennyi *izé* toldalékos formát tartalmazott, például *a mi izénk vőfélyünk kérte ki; olyan kis izével szoktuk megírni olyan kis faggyúval; ha nagyon rossz a tőke akkor csak sárszemre izéljük metsszük*. A gondolati beszédtervezési nehézséget jelzi a következő példa: *hát izé mi mi volt annak a baja, aztán izé /szünet/ mink meg odamentünk*.

5.2. Az *izé* a BEA-adatbázisban

Ennek az adatbázisnak a fejlesztése 2007-ben indult az MTA Nyelvtudományi Intézetében azzal a céllal, hogy jó minőségű hangfelvételeket rögzítsenek különféle, de elsősorban fonetikai elemzések végzésére (Gósy & Gyarmathy, 2017). A BEA különböző típusú spontán beszédanyagokat, mondatismélteléseket és felolvasásokat tartalmaz, a hanganyagok közel 370 órát tesznek ki. A több mint 450 (anonimizált) adatközlő egynyelvű, köznyelvet beszélő budapesti nő és férfi, átlagéletkoruk 36,8 év (18 évestől 90 évesig).

Ebből az adatbázisból véletlenszerűen választottuk ki 50 beszélő (27 nő és 23 férfi) spontánbeszéd-anyagát a 2009-es és 2010-es felvételekből. A Hegedűs Archivum rögzítése óta tehát mintegy 60 év telt el. A BEA-adatbázisból kiválasztott beszélők életkora 20 év és 75 év között szórta (az átlag 53 év). A vizsgált beszédanyag időtartama 30,1 óra, a szószáma 277 374, egy adatközlő átlagosan mintegy 36 percet beszélt, és átlagosan közel 5540 szót ejtett ki.

Elemeztük az *izé* előfordulását és a beszédtervezési diszharmóniára utaló funkcióját. Az 50 felvétel közül 15-ben nem fordult elő a keresett hangsor, 35 beszédanyagban pedig az előfordulás különböző gyakoriságú volt. A beszélőkre jellemzően 1-3 *izé*-ejtés volt található, néhány adatközlő produkált több előfordulást, egy adatközlőnél pedig 58-szor jelent meg. Összesen 162 *izét* azonosítottunk. Az időtartamra vetítve, a keresett hangsor 11,1 percenként fordult elő (5,4 előfordulás óránként); a szószámot tekintve pedig 1712 szavanként ejtettek egy *izé*-t (beszélőre átlagolva ez azt jelenti, hogy mintegy 154 szóból 4,6 volt az *izé*). Gyarmathy ugyancsak a BEA-adatbázis felhasználásával, 35 beszélőnél 141 *izé*-t talált, (2016), ezeket a funkcióik tekintetében és akusztikai-fonetikai szempontból is vizsgálta.

Az anyagunkban talált 162 előfordulást négy fő kategóriába soroltuk, kategorizálásunk kissé eltért az *izé*-vel foglalkozó korábbi két tanulmányétól (vö. Fabulya, 2007; Gyarmathy, 2016). Az általunk kialakított négy fő kategória a következő (példákkal): (i) szótalálási nehézség (pl. *ott kuporgott egy ilyen kis izé*

sánta kiscica), (ii) diszharmónia, amelyet a gondolatnak vagy a megfelelő nyelvi formának a megtalálása okoz (pl. *az a konzi utolsó éve lesz tehát izé akkor ugye az a leghajtósabb év*), (iii) az *izé* tudatos, szándékos használata, amikor nem beszédtervezési nehézség miatt jelenik meg (*tudod, ez egy nagy izé*), valamint (iv) nem egyértelműen kategorizálható előfordulás (pl. *emlékszem nagyszüleimnek olyan volt és de tök nagy volt a csomagtartója /szünet/ így föl lehetett izé és akkor minden belefért*). Vagy a jármű megnevezésére szolgáló lexéma okozott a beszélőnek problémát, vagy a csomagtartó jellemzésére megfogalmazandó gondolat. Az *izé* a kimondandó lemmához tartozó toldalékot is felvette a következő példákban: *így böködte a izének a vakondnak a fejét, szerintem egy izéből vadházasságból ne szülessen gyermek, itt is az van hogy izéenként változik tehát így munkahelyenként változó, jaj most izélsz alszol*.

Anyagunkban 98 előfordulás (60,5%) a szótalálási nehézség kategóriába volt sorolható, 51 példa (31,5%) adódott a gondolati/grammatikai diszharmónia eseteire, 4 adatot (2,5%) találtunk az *izé* szándékos, tehát nem diszharmóniára utaló ejtésére és 9-et (5,5%) a nem egyértelműen kategorizálható előfordulásokra.

Elemzéseink alapján a négy fő kategórián belül további alkategóriák is megkülönböztethetők voltak. A szótalálási nehézség a felszínen homogén kategóriának látszik, az adatok azonban összetettebb képet mutatnak. Szótalálási nehézség áll elő (i) amikor a beszélő nem képes kimondani a kívánt szót, eljut a lemma szintre, de a lexéma előhívásához időre van szüksége (*az eszkimók kunyhójában van izé hidegebb ennél, gyorsan ülj az izé az internet elé, milyen jó az izéje a dolgozata, és akkor lefulladt és ott izé dudáltak meg mutogattak, nem akadt ki az izé a mutató /ti. a sebességmérőé/*), (ii) amikor a beszélő valójában nem is tudja, hogy milyen szót kívánna aktiválni, az is lehetséges, hogy a megfelelő szó nem is tagja a mentális lexikonának (*itt van hogy amit lehet még azt is gyorsan elizéljük /ti. feléljük az energiaforrást/, látsz ilyen embereket akik az izé hogy így az üres fejű búza akinek nincsen sok búra a fejébe, egy ilyen hosszú ház tudjátok egy ilyen izé és akkor több rész van benne, a honvédség a honvédségi izé /a tábor szót kereste, de nem találta/, én hibátlanra csináltam meg a KRESZ izét mindent /a keresett szó lehetett pl. tesztet/, nyilván hát nem ilyen vas izékkal mert a kocsit előbb-utóbb gallyra vágja /?/*), továbbá (iii) amikor versengő helyzet jön létre, a beszélő nem tud hirtelen választani az aktiválható lexémák közül (pl. *az a kicsinek az izéje /szünet/ autója ööö ajándéka, ha egy pohár bor után hogyha valaki izé annyi az al véralkoholfoka vagy alkoholszintje*).

A szótalálási nehézség esetén ejtett *izé*, illetve toldalékolt formái az esetek jelentős részében megfelelő időt biztosítanak a mentális lexikonban történő keresésre és a sikeres találatra. Nemegyszer azonban ez az idő mégsem elegendő, a beszélő nem tudja kiejteni a szükséges szót, ekkor változtat a stratégián. Vagy kimondja a következő szót/szavakat és közbeszúrja a kívánt lexémát (*amikor lépcsőházajtót nyitom ki hogy izé hogy gyakorolsz, voltak olyan izék olyan beugrók, az őserdőbe se szülnek orvossal a izék se szülnek orvossal a régi emberek, hát te egy izé vagy /szünet/ állampolgár vagy, annyira macerás mindig*

így nyári izéket megszerv nyaralásokat megszervezni), vagy újrafogalmazza a közlést. Ez utóbbi történhet egy másik, hasonló jelentésű szó aktiválásával és kiejtésével (pl. *bemész a Csillagpark című /szünet/ izé üzletház vagy nem tudom mi ez ilyen plázaserű, és a túlsó végénél ilyen jó magas ööö meredek izé ilyen dombszerű volt*), avagy a közlés átformázásával (pl. *hát rendesen izé am am amit kért azt megadtuk, hát az egész nap abba a büdös izé fujj, az valami szörnyű*). Egy érdekes példa állt elő, amikor a beszélő a kívánt szó megtalálásával küzdött, de az időnyerő *izé* a megelőző szó toldalékát felvette, mintegy toldalék-perszeveráció történt: *én nem iszom és vezetek a friss jogsival /szünet/ izével /szünet/ jeges úton* (a szándékolt közlés: *én nem iszom és vezetek a friss jogsival jeges úton* volt, az *izé* tehát vagy „*izé-s úton*” vagy „*jeges izé-n*” formában kellett volna, hogy megjelenjen attól függően, hogy melyik szó előhívása okozta a beszélőnek a nehézséget).

A gondolati vagy nyelvi megformálás problémája azt jelenti, hogy a beszélőnek a gondolat kiválasztásával vagy a megfelelő grammatikai formába öntésével állt elő nehézsége. A *meg van egy ilyen izé hogy ööö legyen otthonszülés* példában a beszélő tudja, hogy mi a gondolata, amit közölni akar, de a pontos nyelvi megformáláshoz időre van szüksége, ezért ejti az *izé*-t először, majd e példában hangos szünet is megelőzi a folytatást. Hasonlóképpen, a *hát izé hogy hogy van-e önnek tartós párkapcsolata* példa is a nyelvi megfogalmazás bizonytalanságára utal, vélhetően a téma bizalmas jellege miatt. Az *éjfél előtt jutott eszembe hogy izé elfelejtettem megcsinálni a feladatot* példában feltehetően a gondolat megfogalmazása miatt volt szüksége a beszélőnek többletidőre. Valószínűsíthetően szintén a gondolat kiválasztása eredményezhette az *izé* megjelenését ebben a példában: *nézzétek meg mondjuk a békává ellenőröket azok is izé hogyha véletlenül mondjuk nincsen ráírva a bérletemre az igazolványszámom*. Annak elkülönítése, hogy gondolatszintű vagy grammatikai szintű-e a diszharmónia, nem mindig egyértelmű. A hosszabb közlésrészlet elemzése (a tartalom, a szóhasználat, a grammatikai megformáltság, a megakadásjelenségek, a szünetezés és egyéb jelenségek révén) azonban lehetővé teszi a döntést.

Az *izé* szándékos használata is több okra vezethető vissza, egy szó kiejtésének a tudatos elkerülése, avagy egyfajta rövidítés, mikor a beszélő úgy véli, hogy a beszédpartner az elhangzottak alapján megérti a közlést dacára a „fel nem oldott” *izé* megjelenésének. A *nem akarom minősíteni, de ez egy izé* és a *mert ő hogy kiabált meg hogy milyen izé hogy tehát még ő ő volt negatív* példákban a beszélők trágár, avagy nem szalonképes szavakat kerültek ki az *izé* használatával. A *merthogy merthogy ő tévés meg izé /beszélőváltás következett/* és a *minden nő nagyon szép azokba az izé ruhákba* példákban a beszélők feltételezték, hogy a beszélgetés korábbi részének tartalma alapján a beszélgetőtárs tudja, hogy mi lehet az *izé* „jelentése”, mire utal vele a beszélő (az *izé* egyfajta anaforikus szerepben). Ezek az előfordulások bizonyos kommunikációs konszenzust

feltételeznek a beszédpartnerek között, hiszen a konkrét szavak hiányában is egyértelmű a közlés tartalma.

A rögzített szövegek alapján előfordul, hogy nem egyértelmű, hogy a beszélőnek milyen jellegű diszharmonia okozta a problémát, vagyis az *izé* ejtése szótalálási nehézségre, avagy más beszédtervezési gondra utal. Ilyenek a következő közlések: *jó hát akkor nem mehet tovább és hogy izé rendőrt fognak hívni*. Kérdés, hogy a *rendőr* szó előhívásához, avagy a gondolat kiválasztásának eldöntéséhez kellett a többletidő. A *szerintem izé kínozták az állatokat* példában nyilvánvalónak látszik, hogy a beszélő a *kínoz* szót kereste, ugyanakkor az is lehetséges, hogy a soron következő gondolat megfogalmazásának stilisztikai szempontjai akasztották meg.

6. Következtetések

Az *izé* paradoxona az, hogy noha egy lexikai jelentés nélküli „szó”, a beszélők mégis megértik, felismerik a funkcióját és használják, gyakorlatilag már gyermekkortól. Az *izé* úgy viselkedik, mint egy nyelvi kaméleon (vagy inkább jolly joker), amely bármilyen beszédtervezési diszharmonia esetén gyorsan aktiválható, és rendszerint elegendő időt biztosít a beszélőnek, hogy túljusson a nehézségen. Funkciói kategorizálhatók, ezek eltérő mértékben fordulnak elő a nyelvhasználatban. Bármilyen szófaj helyett állhat, és hordozhatja a megfelelő toldalékot, toldalékokat, ennél fogva egyedülálló jelenség az elemzett nyelvek tekintetében. Némiképpen hasonló a *zera* a baszkban, ugyancsak toldalékolható, azonban jelentős különbség, hogy a *zera* valódi szó, önálló jelentéssel bír (Urizar & Samuel, 2014). Megmagyaráztuk, hogy az *izé* miért nem szó, amikor a spontán beszéd diszharmonias jelenségeinek feloldására használják. Az *izé*-hez azonban kapcsolódhat egyrészt pragmatikai jelentés, másrészt lexikális tartalom (főként főnévi és igei funkcióban): ’nemi szerv’, illetve ’szexuális együttlét’. Egyes kontextusokban a ’vacakol’, ’időt húz’ (*Mit izélsz már annyit!*) jelentésben is használatos. Ezekben az esetekben az *izé* valóságos szó. Előfordulásuk a mindennapi, főként informális beszélgetésekhez köthető (formális narratívákban nem vagy csak nagyon ritkán jelennek meg).

Az *izé* eredete tisztázatlan, a szótárakban többféle elképzelés is fellelhető, amelyeknek azonban egyike sem meggyőző. A jelen tanulmány kísérletet tesz egy teljesen új elméletre az *izé* kialakulásával kapcsolatban, amelyben a nyelvhasználatból, funkcionális elemzésekből indul ki, tekintetbe veszi az artikulációs és beszédpercepciók működéseket, és ilyen módon igyekszik megmagyarázni a fonetikai forma átalakulásait a határozott névelői szerepet kapott névmásoktól a korán leírt és szótározott három beszédhangból álló formáig (1. ábra).

Az *izé* első előfordulása 1604-re datálódik, Szenczi Molnár Albert szótárában található. Az a tény, hogy a szótár tartalmazza, egyértelműen arra utal, hogy a használata már akkorra elterjedt lehetett. A történeti adatok megerősítik, hogy az

izé az elkövetkező évszázadokban aktív tagja volt a nyelvhasználatnak, életkortól, nemtől, műfajtól és nyelvjárástól függetlenül alkalmazták.

Mit mutatnak az adatbázisok a későbbi előfordulásokról? A Hegedűs Archívum és a BEA – amelyek rögzítése között mintegy hatvan év telt el – beszédfelvételeinek elemzése azt sugallja, hogy az *izé* használata kissé növekedhetett a 21. század elején, bár a szakszerű összehasonlítás, számos módszertani okból, nem elvégezhető (a BEA adatbázis felvételi körülményei értelemszerűen befolyásolták a valós nyelvhasználatot, azaz a stigmatizáltságból fakadó tudatos kerülést). Még kevésbé jelenthető ki, hogy a kapott adatok a magyar beszélőközösségekre általánosan jellemzők.

Mi valószínűsíthető tehát az *izé* jelen használatának alakulásáról? Annyi a mindennapi tapasztalat alapján is megállapítható, hogy jóval ritkábban fordul elő, mint akár a kitöltött szünetek (hezitálás), akár a töltelékszók (Libárdi & Gósy, 2020). A használatával kapcsolatban azonban a negatívumok mellett számos pozitívum is felsorakoztatható. Stigmatizációhoz vezet, hogy a két szótagból álló *izé* ejtése meglehetősen feltűnő a verbális kommunikációban, és azonnal informálja a beszélgetőpartnert/partnereket a beszélő tervezési és/vagy kivitelezési problémájáról. Ez nyilvánvalóan zavar(hat)ja a beszélőt. Kevésbé szembevető időnyerő stratégiák is vannak, mint a hezitálás, a valós szavak töltelékszói alkalmazásai, például *hát, tehát, így*, az ismétlések vagy a nyújtások. Ebben a sokféleségben az ismétlődő *izé*-k kirívóak. Tudjuk, hogy az *izé*-nek mind a főnévi, mind az igei formájához társult bizonyos pejoratív, de legalábbis kerülendő jelentéstartalom (a 'nemi szerv', illetve 'szexuális együttlét' értelmezés már a Ballagi-szótárban, illetve Gyarmathinál is szerepel). Ezek a jelentések aktiválód(hat)nak a beszélgetés során, függetlenül attól, hogy a beszélő nem ilyen szemantikai tartalommal ejtette ki az *izé*-t. Használatának kerülésére figyelmeztetik a tanulókat az iskolában.

Ezek után talán meglepő, de az *izé*-nek vannak pozitív tulajdonságai is. Míg a hasonló funkciójú töltelékszókra jellemző a rövidítés a spontán közlésekben, például *szóval – szal, tehát – tát, akkor – akkó*, addig az *izé* évszázadok óta stabilan tartja a három beszédhangot és a két szótagból álló struktúrát. Az időtartamai utalnak a használati funkcióra; szótalálási nehézség esetén az átlagos időtartam 260 ms-nak, a gondolkodási diszharmonia esetén ejtett *izé* átlagos ideje pedig 296 ms-nak adódott (Gyarmathy, 2016). Az *izé* az egyetlen olyan eszköz, amely azonosan működik mind a beszélőben, mind a hallgatóban (ez nem feltétlenül mondható el más időnyerő stratégiáról), és ezáltal funkcionálisan igen jól teljesíti az alkalmazási feladatokat. A szófaji alkalmazkodása és a toldalékolási lehetősége elősegítik a hallgató számára a kiejteni szándékozott lemmát. Mindez pedig nagy kommunikációs előny. A tagadhatatlan pozitívumok a stabilitás és a folyamatos jelenlét zálogait jelenthetik.

Bár egyetértünk az *izé* használatának felsorolt negatívumaival és az ezekből következő stigmatizációval, az mégiscsak tény, hogy a nyelvünkben létezik ez a különleges nyelvi elem, amely oly hosszú időn át ellenállt a törlődésnek, az

emberi felejtésnek. Ez pedig egyértelműen alátámasztja azt, hogy a beszélőknek olykor-olykor szükségük van rá napjainkban is. Egyelőre legalábbis nem tűntethető el a nyelvhasználatból.

Úgy véljük, hogy az *izé* még sok-sok évtizedig lesz tagja a magyar anyanyelvűek mentális lexikonának és része a kollokvialis, kevésbé formalizált, bizalmasabb beszédhelyzeteknek, őrizve ezzel egy – talán nem túlzottan jól hangzó és stigmatizált – speciális „töltelékszót”, de amely egészen sajátos lehetőséget nyújt a beszédtervezési diszharmóniák feloldására (nincs jelentése, könnyen ejthető, morfológiailag bármihez illeszthető), és éppen ezek miatt egyedülálló a világ e tekintetben eddig elemzett valamennyi nyelvéhez képest.

Talán túlzás, de egyfajta hungarikum.

Irodalom

- Anon. szerző (1838). *Magyar és német zsebszótár*. Buda: A Magyar Tudós Társaság.
- Balassa József (1940). *A magyar nyelv szótára 1–2*. Budapest: Grill Károly Könyvkiadóvállalata.
- Ball. = Ballagi Mór (szerk.). (1872). *A magyar nyelv teljes szótára*. Budapest: Franklin Társulat.
- Benkóczy Emil (1908). Népies orvosságnevek. *Magyar Nyelvőr*, 37(4), 188–189.
- Brassai Sámuel (1864). A magyar mondatról III. *Akadémiai értesítő*, 3, 173–409.
- Budenz József (1862). Csuvas közlések és tanulmányok. *Nyelvtudományi Közlemények*, 1, 200–268.
- Czeglédi István (1669). *Redivivus Japhetke*. Kassa.
- CzF. 1 = Czuczor Gergely és Fogarasi János (szerk.) (1862). *A magyar nyelv szótára 1*. Pest: Emich Gusztáv.
- CzF. 3 = Czuczor Gergely és Fogarasi János (szerk.) (1865). *A magyar nyelv szótára 3*. Pest: Emich Gusztáv.
- Dér Csilla Ilona (2020). *Diskurzusjelölők és társulásaik a magyar nyelvben*. Budapest: Károli Gáspár Református Egyetem, L'Harmattan Kiadó.
- ESz. = Zaicz Gábor (főszerk.) (2006). *Etimológiai szótár*. Budapest: Tinta Könyvkiadó.
- EtSz. = Gombocz Zoltán és Melich János (szerk.) (1914–1944). *Magyar etimológiai szótár 1–17. füzet*. Budapest: MTA.
- EWUng. = Benkő Loránd (főszerk.) (1993–1995). *Etymologisches Wörterbuch des Ungarischen 1–2*. Budapest: Akadémiai Kiadó.
- Fabulya Márta (2007). Izé, hogyhívják, hogymondjam. Javítást kezdeményező lexikális kötélemek. *Magyar Nyelvőr*, 131(3), 324–342.
- Fogarasi János (1860). *Német és magyar szótár*. 4. kiadás. *Wörterbuch Ungarischen und Deutschen Sprache* 4. Auflage. Pest: Heckenast Gusztáv.
- Geleji Katona István (1645). *Magyar Grammatikatska*. Gyulafejérvár.
- Gerstner Károly (2016). Lexikatörténet. In Kiss Jenő & Pusztai Ferenc (szerk.), *A magyar nyelvtörténet kézikönyve* (247–308). Budapest: Tinta Könyvkiadó.
- Gósy Mária (1989). *Beszédészlelés*. Budapest: MTA Nyelvtudományi Intézet.
- Gósy Mária, Horváth Viktória & Nikléczy Péter (2011). A Hegedüs-archívum mint korszerű adatbázis. In Bárány M. János & Vargha Fruzsina Sára (szerk.), *Hangok – helyek* (85–103). Budapest: ELTE Magyar Nyelvtudományi és Finnugor Intézet.
- Gósy Mária & Gyarmathy Dorottya (2017). A BEA beszélt nyelvi adatbázis tízéves fejlesztése és kutatási eredményei. *Magyar Nyelvőr*, 141(4), 466–477.
- Gósy, M. & Eklund, R. (2018). Language-specific patterns of segment prolongation in Hungarian. *The Phonetician*, 115, 36–52.
- Gyarmathi Sámuel (1794). *Okoskodva Tanító Magyar Nyelvmester 1–2*. Kolozsvár–Szeben: Hochmeister Márton.
- Gyarmathi Sámuel (1999 [1799]). *Affinitas. A magyar nyelv grammatikailag bizonyított rokonsága a finn eredetű nyelvekkel*. Ford. Koszorús István & Vladár Zsuzsa. Szerk. Szij Enikő. Budapest: Tinta Könyvkiadó.

- Gyarmathy, D. (2016). Izé 'stuff': a corpus-based study on a curious Hungarian filler. *The Phonetician*, 113, 82–101.
- Hamilton, W. L., Leskovec, J. & Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. In Erk, K. & Smith, N. A. (eds.), *Proc. of the 54th Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers)* (1489–1501). Berlin: Association for Computational Linguistics.
- Haspelmath, M. (2023). Defining the word. *Word*, 69(3), 283–297. doi.org/10.1080/00437956.2023.2237272
- Jescheniak, J. D. & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology, Learning, Memory and Cognition* 20, 824–843.
- Juhász Jenő (1959). Izé. Szó- és szólásmagyarázatok *Magyar Nyelv*, 55, 523. https://adt.arcanum.com/hu/view/MagyarNyelv_1959/?pg=530&layout=s&query=iz%C3%A9
- K. Fábián Ilona (1986). Szövegmutatvány Kisszékely nyelvjárásából. *Magyar Nyelvőr* 110(1), 312–319.
- Kiss Jenő 2016. A nyelvi változás. In Kiss Jenő & Pusztai Ferenc (szerk.), *A magyar nyelvtörténet kézikönyve* (20–29). Budapest: Tinta Könyvkiadó.
- Középmagyar emlékirat- és drámakorpusz (<https://ked.nytud.hu/#open>)
- Kresznerics Ferenc (1831). *Magyar szótár gyökérrenddel és deákozattal* 1–2. Buda: Egyetemi Nyomda.
- Kunos Ignác (1887). Orta ojun, török népszínjáték. A varázsló. *Nyelvtudományi Közlemények*, 21(3), 171–190.
- Ladányi Mária (2017). Szavak, szóalakok és szóelemek. In Tolcsvai Gábor (szerk.), *Nyelvtan* (504–515). Budapest: Osiris.
- Laserna, Ch. M., Seih, Yi-Tai, & Pennebaker, J. W. (2014). Um . . . Who like says you know: Filler word use as a function of age, gender, and personality. *Journal of Language and Social Psychology*, 33(3), 328–338. doi.org/10.1177/0261927X14526993
- Levelt, W. J. M. (1989). *Speaking. From intention to articulation*. Cambridge, Mass.: MIT Press. doi:10.5860/choice.27-1947
- Li, Z. & Siew, C. S. Q. (2022). Diachronic semantic change in language is constrained by how people use and learn language. *Memory & Cognition*, 50, 1284–1298. doi.org/10.3758/s13421-022-01331-0
- Libárdi Péter & Gósy Mária (2020). A frázispozíció pszichopragmatikai jelzései tizenévesek beszédében. In Bóna Judit & Krepsz Valéria (szerk.), *Nyelvfejlődés csecsemőkortól kamaszkorig* (107–120). Budapest: ELTE Eötvös Kiadó.
- Loos József (szerk.). (1869–1871). *A magyar, német és tót nyelv szótára. I. kötet: Magyar–német–tót rész. III. kötet: Tót–magyar–német rész*. Pest: Lauffer Vilmos.
- Magyar Történeti Szövegtár (<https://clara.nytud.hu/mts/>)
- Márton József (1807). *Német–magyar és magyar–német lexicon, vagyis szókönyv. Ungarisch–Deutsch und Deutsch–Ungarisch Wörterbuch*. Wien: Haykul Antal.
- Ómagyar korpusz (<http://omagyarkorpusz.nytud.hu/hu-search.html>)
- Pagel, M., Atkinson, Q. D. & Meade, A. (2007). Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature*, 449(7163), 717–720.
- Pápai Páriz–Bod: Dict. = Pápai Páriz Ferenc és Bod Péter (szerk.). (1767). *Dictionarium Latino–Hungaricum. Addita est lingua Germanica*. Szeben: Sárdi Sámuel.
- Podlesskaya, V. (2010). Parameters for typological variation of placeholders. In Amiridze, N., Davis, B. & MacLagan, M. (eds.), *Fillers, pauses and placeholders. Typological studies in language*, 93. (11–32). John Benjamins Publishing Company. (doi.org/10.1075/tsl.93.02pod)
- Prohászka János (1957). Voltaképpen mi is az az izé? *Magyar Nyelvőr*, 81(2), 245–246.
- Schirm Anita (2021). *Diskurzusjelölők szövegeken innen és túl*. Budapest: Loisir Kiadó.
- Simonyi Zsigmond (1876). Egy deminutiv „j” képző nyomai a magyarban. *Magyar Nyelvőr*, 5, 481–484.
- Szathmári István (2007). Mennyiben szolgálták Szenczi Molnár Albert szótárai a magyar irodalmi nyelv (sztenderd) létrejöttét? *Magyar Nyelvőr*, 131(2), 163–172.
- Szenczi Molnár Albert (1604). *Dictionarium Latinoungaricum et Dictionarium Ungaricolatinum*. Nürnberg: Elias Hutter.

- Szinnyei József (1911). Adalékok a magyar nyelv szófejtő szótárához. *Nyelvtudományi Közlemények*, 41, 110–112.
- SzófSz. = Bárczi Géza (szerk.). (1941). *Magyar szófejtő szótár*. Budapest: Királyi Magyar Egyetemi Nyomda.
- Taylor, J. R. (ed.) (2014). *The Oxford handbook of the word*. Oxford: Oxford University Press. doi.org/10.1093/oxfordhb/9780199641604.001.0001
- TESz. = Benkő Loránd (főszerk.) (1967–1976). *A magyar nyelv történeti-etimológiai szótára 1–3*. Budapest: Akadémiai Kiadó.
- Történeti Magánéleti Korpusz (<https://nytud.hu/adatforras/o-es-kozepmagyar-torteneti-maganeleti-korpusz>)
- ÚESzWeb. = Gerstner Károly (szerk.). (2023). *Új magyar etimológiai szótár*. <https://uesz.nytud.hu/index.html>
- Urizar, X. & Samuel, A. G. (2014). A corpus-based study of fillers among native Basque speakers and the role of *zera*. *Language and Speech*, 57(3), 338–366.
- Varga Ignác (1908). Népnyelvi hagyományok. Elbeszélés. *Magyar Nyelvőr*, 37(1), 89–90.
- Vass József (1864). Adalékok a magyar nyelvjárástanhoz. *Nyelvtudományi Közlemények*, 3(1), 1–18.
- VégSz. = Papp Ferenc (szerk.). (1969). *A magyar nyelv szóvégmutató szótára*. Budapest: Akadémiai Kiadó.
- Wang, B., Di Buccio, E. & Melucci, M. (2021). Sequential modeling in vector space. *Proceedings of the 11th Italian Information Retrieval Workshop 2021, Bari, Italy, September 13-15*. <https://ceur-ws.org/Vol-2947/paper12.pdf>

DODÉ RÉKA – YANG ZIJIAN GYŐZŐ

HUN-REN Nyelvtudományi Kutatóközpont
dode.reka@nytud.hun-ren.hu
<https://orcid.org/0009-0008-3880-2750>

yang.zijian.gyozo@nytud.hun-ren.hu
<https://orcid.org/0000-0001-9955-860X>

Dodé Réka – Yang Zijian Győző: Generálás vs. kinyerés: nagy nyelvi modellel kapott kulcsszavak vizsgálata

Alkalmazott Nyelvtudomány, XXV. évfolyam, 2025/1. szám, 67–83.
[doi:http://dx.doi.org/10.18460/ANY.2025.1.004](http://dx.doi.org/10.18460/ANY.2025.1.004)

Generálás vs. kinyerés: nagy nyelvi modellel kapott kulcsszavak vizsgálata

Two types of keyword extraction methods can be distinguished: extractive and abstractive. Extractive keyword extraction relies on author-provided keywords, selecting the most appropriate ones for the given text. Abstractive keyword generation, on the other hand, is more similar to how a human would generate keywords. This method can generate keywords that may not have appeared in the author-provided list.

In our research, we employ the abstractive method to generate keywords by fine-tuning the PULI Llumix 32K language model, which processes inputs of up to 32,000 tokens. We examine whether the keywords generated by the language model match the author-provided keywords and whether they appear in the text. If they differ from the author’s keywords, we analyze the extent and nature of these differences—whether they are more general, more specific, synonyms, or other variations—evaluate their relevance to the text’s topic, and determine if they accurately represent the text.

The purpose of keywords is to represent the text, provide insight into its content, and facilitate easier navigation among technical texts. However, the strategy behind author-provided keywords differs from a simple frequency-based approach.

The fine-tuning of the PULI Llumix 32K language model, with a 32,000-token input, was based on author-provided keywords extracted from studies in the REAL repository. We briefly present the training data and process used for fine-tuning based on an already published study.

We further examined 133 manually reviewed results. The resulting keywords were classified into four groups: 1) entirely new (neither in the text nor among the keywords), 2) learned (among the keywords but not in the text), 3) extracted (not among the keywords but present in the text), and 4) accurate/matched (present both in the keywords and the text). In the first part of the research, we quantify these groups. We then take a closer look at the first group and the third group: how they relate to the author-provided keywords (more general, more specific, synonyms, or other differences), their relevance to the text’s topic, and whether they accurately represent the text.

The analysis indicates that the language model effectively generates an appropriate number of keywords, with 44.2% matching those provided by authors. Additionally, 21.8% of the model-generated keywords do not appear directly in the text, suggesting creative inference. However, categorization challenges arise, particularly with synonyms and domain-specific terms, due to the model’s limited conceptual understanding.

Compared to author-assigned keywords, the model’s generated keywords differ in distribution, often leaning toward more general concepts. While some keywords align with indexing principles, others contradict specificity guidelines by being overly broad. Multi-word keywords, which are more thematically relevant, mirror patterns observed in human-generated keywords, making them more effective in representing text content.

Zero-shot prompting yields inconsistencies, resulting in generic keywords that weaken both search engine efficiency and text representation. Fine-tuning the model with additional data or modifying the prompting approach could enhance its performance, encouraging more precise and contextually relevant keyword generation.

This research contributes to the testing and refinement of keyword generation models, offering insights for improving future fine-tuning strategies. Adopting reinforcement learning based on human feedback could further optimize the model, fostering the generation of creative, multi-word, and highly specific keywords.

Keywords: PULI Llumix 32K LLM, fine-tuning, zero-shot prompting, generated keyword, extractive keyword extraction

1. Bevezetés

Aki a tudományos életben tevékenykedik és publikál, jól ismeri a kulcsszavakat és azok célját. Számos területen bevett gyakorlat, hogy a szerzők kulcsszavakat adnak meg tudományos publikációikban, hogy azok segítsék az olvasókat. A kulcsszavak több, fontos funkciót töltenek be: betekintést nyújtanak a szöveg tartalmába, képviselik a fő témákat, megkönnyítik az akadémiai szövegek közötti navigációt, emellett növelik a láthatóságot a keresőmotorok által. Adódik azonban a kérdés, hogy milyen kulcsszavakat érdemes ehhez megadni?

Az UNESCO (1975) meghatározta azokat az alapelveket, amelyeket a kulcsszavaknak követniük kell, függetlenül attól, hogy manuálisan vagy automatikusan kerültek-e meghatározásra. A két fő alapelv a teljesség és a specifikusság. Firoozeh és munkatársai (2020) pedig további elvekkel egészítik ki ezeket: minimalitás, pártatlanság és reprezentativitás.

A teljesség azt jelenti, hogy a kulcsszavaknak le kell fedniük a dokumentum minden olyan témáját, amely potenciális információs értékkel bír, beleértve az implicit, a dokumentumban nem kifejezetten megjelenő témákat is. A specifikusság elve szerint olyan nyelvi elemeket kell kiválasztani, amelyek kifejezetten az adott dokumentumra jellemzők, és nem alkalmazhatók általánosan hasonló dokumentumokra (UNESCO, 1975). Firoozeh és társai (2020) szerint a minimalitás elve megköveteli, hogy a kulcsszavak egyedi jelentésűek legyenek – ez növeli a konzisztenciát és az érthetőséget is. A pártatlanság elve az objektivitást hangsúlyozza, míg a reprezentativitás elve szerint a kulcsszavaknak a dokumentum lényeges elemeit kell tükrözniük.

Dodé (2024) egyik tanulmányában a szerzői kulcsszavak szövegbeli előfordulását vizsgálta. Rámutatott, hogy az általa vizsgált korpuszban a kulcsszavak 37%-a kevesebb mint kétszer fordul elő a szövegekben. Ennek oka azonban nem csupán elírások vagy morfoszintaktikai különbségek, hanem az is, hogy a szerzők szándékosan általánosabb vagy specifikusabb kulcsszavakat választanak, figyelembe véve azt, hogy a szöveget a szakmai közönség olvassa majd.

A kulcsszóadás tehát egy átgondolt stratégia, amelyet automatikus kinyeréssel vagy épp generálással próbálunk reprodukálni, automatizálni.

A kulcsszókinyerés alapvetően egy szövegelemzési technika, amelyet a természetesnyelv-feldolgozás, az információkeresés és a szövegbányászat alkalmazásaiban használnak (Firoozeh et al., 2020). A kulcsszókinyerés hosszú múltra tekint vissza. Namoto (2023) az elmúlt 50 év kulcsszókinyerési módszereit foglalta össze tanulmányában, amelyben helyet kapnak az olyan módszerek, mint az automatikus terminusindexelés, a nyelvtaniszabály-alapú kinyerés, a gráfalapú rangsorolási modell, a külső információt alkalmazó modell, a szövegosztályozás, a különböző statisztikák, majd a generatív modellek. Nomoto (2023) végül javasolja a generatív rendszerekre való áttérést is, mivel ezek lehetőséget biztosítanak kulcsszavak belső és külső építésére. A generatív rendszerekkel pedig elérkeztünk a neurális nyelvi modellekhez és a generáláshoz.

Egy másik szempontból közelítve, kétféle módszert különböztethetünk meg: extraktív és absztraktív. Extraktív módszernek azt hívjuk, amikor a modellnek meglévő címkék közül kell választania. A fent említett megoldások közül például a terminusindexelés vagy a szövegosztályozás tartozik ebbe a kategóriába. A másik módszer az absztraktív, amikor a modell képes a megadott címkéken kívül teljesen új elemet létrehozni. Az utóbbi években a generatív nyelvi modellek lehetővé tették az absztraktív kulcsszógenerálást.

1.1. A nyelvmodellek

Az elmúlt évtizedben a neurális nyelvi modellek megjelenése alapvető változásokat hozott a nyelvtechnológia területén. Ez eleinte statikus szóbeágyazásokat jelentett, azonban mára számos transzformeralapú és generatív kontextuális nyelvmodell is rendelkezésünkre áll.

A nyelvi modellek a közelgő szavak előrejelzését adják a megelőző szókontextusból (Jurafsky–Martin, 2023). A korábbi nyelvi modellekkel szemben a neurális (nagy) nyelvi modellek „can handle much longer histories, can generalize better over contexts of similar words, and are more accurate at word-prediction” (Jurafsky–Martin, 2023: 147).¹ A *nagy* jelző a betanítóadatok hatalmas mennyiségére utal.

A mély neurális hálózatok elterjedésével lehetőség nyílt a vektoriális szemantika világából ismert szóbeágyazás (word embedding) (Mikolov et al., 2013) módszertanát alkalmazni. Egy beágyazás minden szóhoz egy sokdimenziós, folytonos vektort rendel. Ezek a vektorok egy szemantikus teret feszítenek ki, ahol a hasonló jelentésű szavak vektorai közel esnek egymáshoz. Ezek a modellek a szemantikai jegyek mellett szintaktikai jellemzőket is képesek megtanulni (Nemeskey, 2020).

A kezdeti statikus szóbeágyazáson alapuló nyelvmodelleket idővel felváltották a kontextuális transzformeralapú nyelvmodellek. A transzformer modell (Vaswani et al., 2017) az „önfigyelmi” (self-attention) mechanizmuson alapszik.

¹ „sokkal hosszabb előzményeket tudnak kezelni, jobban általánosíthatók a hasonló szavak kontextusában, és pontosabbak a szó előrejelzésében” (ford.: a szerzők)

A modell képes környezetfüggő szövegreprezentációt képezni, amelynek segítségével pontosabban tudja megtanulni a nyelvi jelenségeket. Az eredeti transzformer modellt egy enkóder és egy dekóder architektúrával tervezték meg, ahol az enkóder rész felel azért, hogy bemeneti szövegből egy vektor reprezentációt készítsen és a dekóder felel a szöveg generálásáért. A Transzformer modellen alapuló modellek megjelenése után szignifikánsan felülmúlta az addigi legjobb eredményeket (Barrault et al., 2020). A transzformer architektúráját tekintve, mind az enkóder, mind a dekóder külön-külön is használhatók és tanítható belőlük nyelvmoddell. Tipikusan egy osztályozó feladatot, mint amilyen a kulcsszókinyerő feladat is lehet egy csak enkóder alapú nyelvmoddellel, például egy BERT-moddellel (Devlin et al., 2019) tanítani. Azonban az enkóderalapú modellekkel nem vagyunk képesek új tartalom generálására, ezért ha a feladat célja az, hogy új kulcsszavakat is tudjon generálni, akkor egy dekóderalapú modellt érdemes választani.

A csak dekóderalapú modellek családjába tartoznak az utóbbi évek legnépszerűbb modelljei; a nagy nyelvi modellek, mint a GPT-modellek (Brown et al., 2020), illetve az azokra épülő alkalmazások, mint a ChatGPT². Az utóbbi hónapok kutatásai egyértelműen azt bizonyítják, hogy ha kellően nagy egy ilyen csak dekóderalapú nyelvi modellt, és kellően sok adattal tanították, akkor képesek rendkívül széles körben megoldani nyelvtechnológiai feladatokat, köztük például a kulcsszókinyerést. Azonban az ilyen nyersen előtanított nagy nyelvi modellt az úgynevezett „következő szó prediktálása” feladattal tanították elő, ahol azt tanulja meg, hogy egy adott bemeneti szöveget hogyan kell folytatni. Így, ahhoz hogy kulcsszókinyerésre vagy kulcsszógenerálásra használhassuk, akkor azt tovább kell finomhangolni az adott feladatra.

Egy ilyen jellegű feladathoz, mint a kulcsszó generálása, többféle nyelvmoddelt lehet választani. Mivel magyar nyelvre szeretnénk készíteni modellt, ezért olyan modellek jöhetnek szóba, amelyek magyarnyelv-tudással rendelkeznek. A nagy nyelvi modellek korszaka előtt több magyar nyelvű modellt is létrejött, mint a huBERT (Nemeskey, 2021), vagy PULI BERT-Large vagy PULI GPT-2 (Yang et al., 2023a), azonban egyrészt ezek a modellek korlátozott bemeneti hosszra képesek kezelni, másrészt az új trendeket követve a nagy nyelvi modellekkel optimálisabb eredményeket lehet elérni.

Jelenleg négy nagy nyelvi modellt érhető el magyar nyelvre. Az egyik a HILANCO GPTX³ (Feldmann et al., 2021), egy angol-magyar kétnyelvű GPT-NeoX típusú modellt (Andonian et al., 2023). Ez sajnos csak korlátozottan elérhető. Emellett rendelkezésre állnak a PULI-család⁴ GPT-NeoX (Antonian et al., 2023) típusú nagy nyelvi modelljei, köztük a magyar nyelvű PULI 3SX (Yang et al., 2023a), valamint a háromnyelvű (magyar-angol-kínai) PULI Trio modellt

² <https://chat.openai.com>

³ <https://hilanco.github.io/home.html>

⁴ <https://puli.nytud.hu/>

(Yang et al., 2023b). Végül a jelen kutatás során használt többnyelvű Llama-2 (Touvron et al., 2023) magyar nyelvre továbbtanított modellje, a PULI LluMiX 32K (Yang et al., 2024).

Az általános nyelvmodellek óriási mennyiségű szöveges adaton tanulnak, és az abból megszerzett tudásból építkeznek, így olyan problémákat is képesek megoldani, amelyeket a régebbi alkalmazások nem vagy nem olyan hatékonyan (pl. utasításkövetés).

1.2. Kapcsolódó irodalom

A kapcsolódó irodalmak között fontos megemlíteni ennek a kutatásnak az alapját képező, ezt a kutatást motiváló kutatást; Dodé (2024) a szerzői kulcsszavakat és azok szövegbeni előfordulását vizsgálta, majd összevetette őket a terminusok tulajdonságaival. Egy másik kutatás során Dodé (2023) különböző promptokkal kulcsszó- és terminuskinyerési kísérleteket végzett a ChatGPT-t segítségével. A prompt egy olyan utasítás, kérdés vagy kérdés, amelyre válaszként kimenetet kapunk a finomhangolt modelltől (Jalalov–Gaszcz, 2023). A promptok fontosságát hangsúlyozza Reynolds és McDonell is a nyelvmodellek hatékony kiaknázása kapcsán (Reynolds–McDonell, 2021).

A promptkutatást követően magyar nyelven továbbtanított nyelvmodellel kísérletezett Dodé és Yang (2024a, Dodé–Yang, 2024b). Kutatásukban statikus szóbeágyazáson alapuló nyelvmodellt tanítottak be a fastText (Joulin et al., 2017) segítségével, valamint finomhangoltak transzformeralapú nagy nyelvmodelleket, mint a PULI 3SX és Llama-2. Egyértelműen alátámasztják azt, hogy a transzformeralapú nyelvmodell szignifikánsan jobb eredményt ér el, mint a statikus szóbeágyazáson alapuló modell, valamint nem elhanyagolható szempont az is, hogy a kutatásban használt cikkek rendkívül hosszú dokumentumok, így az is fontos szempont, hogy olyan Llama-2 modellvariánst alkalmaztak, ami megnövelt bemeneti hosszal rendelkezett.

Továbbá érdemes megemlíteni, hogy a kulcsszógenerálási feladat kétféleképpen is megoldható a nyelvmodellek segítségével. Az egyik megoldás lehet az úgynevezett extraktív megközelítés, amikor az adott szöveg alapján a meglévő címkék közül választ ki a modell releváns kulcsszavakat. A statikus szóbeágyazáson alapuló nyelvmodell csak extraktív módon képes kulcsszókinyerést végezni. A másik megoldás az absztraktív kulcsszógenerálás, amikor a modell képes az adott szöveg alapján az előzőleg megadott kulcsszavakhoz képest teljesen új kulcsszavakat generálni. Ehhez a feladathoz mindenképpen szükséges egy dekéderalapú architektúrájú nyelvmodellt használni. A korábbi kutatásunkban mindkét irányt kipróbáltuk, de egyrészt a nagy nyelvi modellek szignifikánsan jobb eredményt értek el, másrészt szeretnénk, ha a modell képes lenne kreatív, új kulcsszavakat generálni, ezért végül az absztraktív kulcsszógenerálást választottuk végső megoldásnak.

Yang és munkatársai (Yang et al., 2020) az extraktív módszert alkalmazva fejlesztett magyar nyelvű kulcsszó- és címkekinyerő rendszer, amelyet sajtó szövegekkel finomhangoltak. Kutatásukban szintén a fastText implementációt alkalmazták.

Az elmúlt évek során többen is foglalkoztak azzal, hogy a nagy nyelvi modelleket terminológiai feladatokra, például terminuskivonásra és kulcsszókinyésre finomhangolják (például Sun et al., 2020; Sammet–Krestel, 2023, Liu et al., 2021).

2. Módszertan

A kutatáshoz használt anyag egy már meglévő kutatásból származó eredmény további vizsgálata.

2.1. A tanítóanyag és a PULI Llumix 32K nyelvi modell betanítása

Ez a fejezet Dodé–Yang (2024a) tanulmánya alapján készült, annak röviden bemutatott eredménye.

A PULI Llumix 32K modellt a HUN REN Nyelvtudományi Kutatóközpont fejlesztette (Yang et al., 2024). A modell egy 32 ezres kontextus ablakú, magyar nyelvre továbbtanított Llama-2 modell. Az eredeti Llama-2 modellt a Together⁵ továbbtanította, és bemeneti hosszát a finomhangolás során megnövelték 32.768 hosszúra (Llama-2-7B-32K). Ezzel lehetőség nyílt hosszú dokumentumok feldolgozására. Ezután Yang és munkatársai (2024) ezt a megnövelt bemenetű Llama-2-7B-32K modellt tanították tovább előtanítási módszerrel nyers magyar korpuszon. Ezzel gyakorlatilag magyar nyelvre adaptáltak egy Llama-2 modellt. A tovább előtanításhoz körülbelül 8 milliárd magyar szót használtak. A továbbtanítás tudásfelejtéssel járhat, ezért a 8 milliárd magyar szó mellé körülbelül 2 milliárd szónyi angol szöveget is keverték. Majd az így elkészült modellt finomhangolták kulcsszógenerálás feladatára.

A finomhangolás folyamata során egy előzetesen betanított modellt továbbképeznek egy adott feladatra (Jurafsky–Martin, 2023). Ilyen feladat például a koreferenciafeloldás. Az előtanítás során a modell hatalmas mennyiségű szöveg feldolgozásával megtanulja a szavak és mondatok jelentésének bizonyos reprezentációit (Jurafsky–Martin, 2023).

A kutatás a REAL-repozitórium (tudományos tanulmányok tára) egy részének felhasználásával készült. A PDF-anyagokat először txt formátumba konvertálták, végül 1146 szövegből nyerték ki a szerző által megadott kulcsszavakat, amivel a PULI Llumix 32K modelljét finomhangolták. Az összehasonlítás kedvéért, az eredeti Llama-2-7B-32K modellt is finomhangolták és egyértelműen kimutatták, hogy a magyarra adaptált változat jobb eredményt ért el, vagyis a továbbtanítás magyar nyelvre egyértelműen értékes magyar tudást adott az eredeti modellhez.

⁵ <https://www.together.ai/>

Így mi a továbbiakban már csak a PULI Llumix 32K modellt fogjuk tanítani és elemezni.

A kutatásukhoz a korpuszt felbontották tanító- és tesztanyagra – 1000 cikket tanításra és 145 cikket tesztelésre. A tanításhoz a Stanford Alpaca promptsablont és az OpenChatKit implementációját használták (lásd az 1. ábrán).

1. ábra. Stanford Alpaca promptsablon és az OpenChatKit implementáció

Stanford Alpaca
Az alábbiakban egy utasítást találsz, amely leír egy feladatot, amelyhez egy bemenetet is mellékelünk, hogy további összefüggéseket adjon. Írj egy választ, amely megfelelően teljesíti a feladatot! (Below you'll find an instruction describing a task, along with an input provided to offer further context. Write a response that adequately fulfills the task!)
Instruct:
Generálj kulcsszavakat az alábbi szöveg alapján! (Generate keywords based on the provided text!)
Input:
[content of the article]
Answer:
[keywords]
OpenChatKit
[content of the article]
<Q>: Generálj kulcsszavakat a megadott szöveg alapján! (Generate keywords based on the provided text!)
<A>: [keywords]

A jelen kutatáshoz azután a 145 cikket tartalmazó tesztalmaz eredményeit használtuk. A 145 tételt végignéztük, és kiszűrtük belőle a biztosan hibásan felismert vagy hibásan konvertált szövegeket (például azok az esetek, ahol a kulcsszavak nem vesszővel, hanem valamilyen más szimbólummal voltak elválasztva, és ezt nem ismerte fel az OCR), így kaptuk meg a végleges 133 db tételt. Ezekhez a következő adatok álltak rendelkezésre (1. táblázat).

1. táblázat. A PULI Llumix 32K modell tesztalmazához tartozó adatok

A referenciakorpusz (a szerző által megadott kulcsszavak)	A PULI Llumix 32K modell eredményei
	A szövegben előforduló kulcsszavak.
	A szövegben előforduló kulcsszavak száma.
	A szövegben nem talált kulcsszavak.
	A szövegben nem talált kulcsszavak száma.

A szerzők által megadott kulcsszavak száma összesen a 133 szövegnél 621 db, míg a PULI Llumix 32K modell eredményeinél 658 db, átlagosan tehát 4,67 kulcsszó volt szövegenként a referenciaanyagban és 4,95 a modell anyagában.

2.2. A kutatás kérdései és a vizsgálat módszerei: a kapott kulcsszavak kategorizálásai

A kutatás során arra voltunk kíváncsiak, hogy a nyelvmodell által kapott kulcsszavak mennyiben újak; mennyiben újak a szöveghez képest – tehát szerepelnek-e a szövegben – és mennyiben újak a szerzői kulcsszavakhoz képest. Kíváncsiak voltunk továbbá, hogy az eltérések milyen típusúak, és hogy relevánsnak tekinthetők-e a szöveg témáját és a szerzői kulcsszavakat tekintve.

Ezeknek a kérdéseknek a megválaszolásához a modell által kapott kulcsszavakat négy csoportba soroltuk:

1. **Teljesen új:** a kapott kulcsszó nincs a szövegben és nem egyezik a szerzői kulcsszóval.
2. **Tanult:** a kapott kulcsszó nincs a szövegben, de megegyezik a szerzői kulcsszóval.
3. **Kinyert:** a kapott kulcsszó szerepel a szövegben, de nem egyezik a szerzői kulcsszóval.
4. **Megegyező:** a kapott kulcsszó szerepel a szövegben és megegyezik a szerzői kulcsszóval.

A következőkben az első és a harmadik (tehát, amelyek eltértek a szerzői kulcsszavaktól) csoport kulcsszavait vizsgáltuk. Ekkor arra voltunk kíváncsiak, hogy az új kulcsszavak milyen típusúak, hogyan térnek el a szerző által adott kulcsszótól.

A kategóriák ez esetben a következők voltak:

1. Eltérő végződés (képző és rag is).
2. Eltérő helyesírás/írásmód (jelölt-jelöletlen szerkezetek is).
3. Fordítás: mindkét irányú.
4. Speciálisabb kulcsszó a szerzőjéhez képest: a kulcsszó kiegészül egy plusz tulajdonsággal, szerkezetként, összetételként szerepel (pl. *kiképzés* – *katonai kiképzés*).
5. Általánosabb kulcsszó a szerzőjéhez képest: az előző folyamat fordítottja (pl. *digitális kompetencia profil* – *digitális kompetencia*).
6. Rövidítés/potenciális szinonima.
7. Eltérő/nem kategorizálható: látszólag nem kapcsolódó, vagy nem közvetlenül kapcsolódó kulcsszavak, illetve asszociatív viszonyban állnak egymással.
8. Szerkezet: mellérendelést tartalmazó szerkezetek, illetve tagmondatjellegű szerkezetek.

A kategóriák után néhány példát részletesebben is megvizsgáltunk, illetve a szerző által adott, a szövegben nem szereplő esetek közül is bemutatunk néhányat.

2.3. Kulcsszavak a fogalmi struktúrában

Kétségtelen, hogy a hierarchikus rendszerezés az, amely a leginkább megfelel a dolgok rendszerré szervezéséhez.

A kulcsszavak tágabb értelemben véve terminusoknak tekinthetők (vö. Dodé, 2024), így vizsgálhatjuk őket a fogalmak rendszerében, ahogy azt a terminológiatudomány teszi. A „fogalmak nem elszigetelt tudásegységként léteznek, hanem mindig egymáshoz viszonyítva” (ISO 704: 2022)⁶.

A fogalmi rendszer asszociatív és hierarchikus kapcsolatokról áll (ISO 704: 2022). A hierarchikusan belül beszélhetünk általános (generikus-specifikus) és partitív (rész-egész) viszonyokról. Egy specifikus fogalom feltételez egy általános vagy általánosabb fogalmat is, amelybe/amelyhez tartozik.

Szövegszinten az általános és a specifikus kapcsolatot lehet tetten érni a kollokációkon keresztül. Egy-egy kulcsszó összes alakjának kollokációira rákeresve, fel tudunk térképezni egy-egy fogalmi részrendszert (Kis–Pohl, 2005: 224).

A fogalmak nemcsak hierarchikus kapcsolatokban állhatnak egymással, hanem asszociatív viszonyokban is. Ez azt jelenti, hogy a fogalmak között szorosabb, de nem hierarchikus kapcsolatok is létrejöhetnek. Ennek számos fajtája van a *Terminology work — Principles and methods* című nemzetközi ISO-szabvány alapján (ISO 704: 2022):

1. sequential relation: szekvenciális kapcsolat – az objektumok valamilyen rendezésén alapul (pl. idő, következmény);
2. activity relation: tevékenységen alapuló kapcsolat (pl. ágens kapcsolat, célkapcsolat);
3. origination relation: az objektum eredetére vonatkozó kapcsolat (eredetkapcsolat, hozzávalókapcsolat);
4. instrumental relation: eszközkapcsolat (objektum-eszköz kapcsolat);
5. interactional relation: kétirányú kapcsolat az objektumok között (függőségi kapcsolat);
6. transmission relation: feladó–címzett elv alapján (feladókapcsolat, fogadókapcsolat);
7. opposite relation: ellentétes kapcsolat (ellentmondásos kapcsolat).

Ezeket a típusú kapcsolatokat azonban a szövegből, szaktudás nélkül, manuális úton nehéz detektálni, illetve ezeknek a kapcsolatoknak, relációknak a feltérképezésére az ontológia hivatott (Szakadát–Szóts–Szaskó, 2006).

2.4. A vizsgálat lehetséges hibái

A kutatásnak vannak bizonyos hibafaktorai. Az egyik a szövegfelismerésből fakad. Bár az eredményeket átnéztük és a kulcsszavaknál detektált hibásan felismert szövegeket kiszűrtük, azonban a tanulmányok szövegében szintén előfordulhatnak felismerési hibák, amik ronthatják a keresés eredményét. Ilyen

⁶ „Concepts do not exist as isolated units of knowledge but always in relation to each other.”

például az elválasztás, vagy akár a nem felismert szimbólumok esete, illetőleg azokban a tanulmányokban, amelyek hasábos elrendezésben jelentek meg, különböző 'feldarabolt' szavak.

A másik hibája, torzítása az eredményeknek abból fakadhat, hogy a szövegbeli előfordulást egy egyszerű kereséssel végeztük – szó szerinti egyezéseket számoltunk, miután a szöveget kisbetűssé alakítottuk –, amely így a módosult végződésű, illetve bizonyos összetételekben előforduló kulcsszavakat valószínűleg nem találta meg.

2.5. Zero-shot prompting GPT-4o mini modellel

Kutatásunkban arra is kíváncsiak voltunk, hogy egy népszerű, sokak által használt nagy nyelvmodell – a GPT-4o mini⁷ – milyen kulcsszavakat generál a szerzői kulcsszavakhoz (ilyen jellegű kutatás: Dodé, 2023) és a PULI LlumiX 32K modellhez képest. Az eredmények összehasonlításánál és értékelésénél fontos megjegyezni, hogy míg a PULI LlumiX 32K modell finomhangolva lett a feladatra, tehát előre megkapott 1000 szöveget és a hozzájuk tartozó szerzői kulcsszavakat, addig a GPT-4o mini modell csupán egy promptot kapott, amiben ugyanazt kértük tőle, mint a PULI LlumiX 32K modelltől:

Instruct: „Generálj kulcsszavakat az alábbi szöveg alapján! (Generate keywords based on the provided text!)”

Imput: [content of the article]

Ennek a technikának a neve a zero-shot prompting. Ez egy olyan promptolási technika, ahol a nyelvmodell nem lett előzetesen betanítva és példát sem lát egy adott feladatra, csupán egy promptra adott válaszként reagál az instrukcióra (Syed–Gadesha, 2025). Ennek fényében tehát az összehasonlítás nem validálható, viszont bizonyos minták kirajzolódnak a finomhangolt nyelvmodell és a zero-shot prompting közti különbség kapcsán.

3. Eredmény

3.1. Az új kulcsszavak kvantitatív elemzése

A 133 szöveghez összesen 621 szerzői kulcsszó tartozott, amelyből 23%, azaz 143 nem szerepelt a szövegben. A nyelvmodell által kapott kulcsszavaknál az eredmény 21,9% volt, azaz a 658 darab kulcsszóból 21,9%, 144 darab nem volt megtalálható a szövegben a keresési módszerünkkel. A módszereknél vázolt négy csoport eredményeit a 2. táblázat tartalmazza.

⁷ <https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/>

2. táblázat. Az új kulcsszavak négy csoport szerinti eredményei

Teljesen új		Tanult		Kinyert		Megegyező	
db	%	db	%	db	%	db	%
122	18,5	22	3,3	245	37,2	269	40,9

A legkevesebb, kicsivel több mint 3% volt az olyan kulcsszavak száma, amelyek a szerző által adott kulcsszavakkal megegyeztek, azonban a szövegben nem szerepeltek. A kinyert és a megegyező kategóriák mennyisége hasonló volt (37,2% és 40,9%), míg a teljesen új is majdnem 20%.

3.2. Az új kulcsszavak kvalitatív elemzése

A teljesen új és a kinyert kategóriákban használt további csoportosítás eredménye a következő táblázatban (3. táblázat) látható.

3. táblázat. A teljesen új és a kinyert kulcsszavak további kategorizálása szerinti eredmények

	Teljesen új (db)	Teljesen új (%)	Kinyert (db)	Kinyert (%)
Eltérő végződés	1	0,8%	12	4,9%
Eltérő (helyes)írás	12	9,8%	17	6,9%
Fordítás: mindkét irányú	6	4,9%	7	2,9%
Speciálisabb kulcsszó a szerzői kulcsszóhoz képest	9	7,4%	14	5,7%
Általánosabb kulcsszó a szerzői kulcsszóhoz képest	5	4,1%	37	15,1%
Rövidítés/potenciál is szinonima	7	5,7%	12	4,9%
Eltérő/nem kategorizálható	67	54,9%	141	57,6%
Szerkezet	15	12,3%	5	2%

A következőkben néhány példát mutatunk be kategóriánként.

4. táblázat. Példák a különböző kategóriák kulcsszavaira

	Referencia-kulcsszó (szerzői)	Eredménykulcsszó (nyelvmodell)
Eltérő végződés	<i>rehidrálás</i>	<i>rehidráció</i>
Eltérő (helyes)írás	<i>magia naturalis</i>	<i>mágia naturalis</i>
	<i>egyetemi integráció</i>	<i>egyetemintegráció</i>
Fordítás: mindkét irányú	<i>családkutatás</i>	<i>familysearch</i>
Speciálisabb kulcsszó a szerzői kulcsszóhoz képest	<i>kiképzés</i>	<i>katonai kiképzés</i>
	<i>egyháztörténet</i>	<i>egyháztörténet-írás</i>
Általánosabb kulcsszó a szerző által adott kulcsszóhoz képest	<i>gravitációs modell jel-kód</i>	<i>gravitációs modell</i>
	<i>infrastruktúra mint szolgáltatás</i>	<i>infrastruktúra</i>
	<i>környezetvédelem</i>	<i>környezet</i>
Rövidítés/potenciál is szinonima	<i>szexualitás/szerelem</i>	<i>szerelem/szexualitás</i>
	<i>büntetés-végrehajtási reintegrációs célok</i>	<i>bv. reintegrációs célok</i>
Eltérő/nem kategorizálható	<i>református mikroközösség</i>	<i>református mikrotörténet</i>
	<i>evangélikus iskoláztatás</i>	<i>evangélikus oktatásügy</i>
	-	<i>*szociokonómiai helyzet</i>
	-	<i>taxonómia</i>
Szerkezet	-	<i>bolygatott 10. századi temető</i>
	-	<i>minőségirányítás az oktatásban</i>

3.3. A GPT-4o mini által generált kulcsszavak

Mivel a GPT-4o mini modell nem lett előre betanítva – és az instrukcióban sem maximalizáltuk a kulcsszavak számát –, így nem tudott rátanulni a kulcsszavak számára, és minden esetben jóval több kulcsszót generált (a maximum 50 db volt). A modell válaszára többféle példát lehetett látni:

1. Bevezetővel kezdi a választ: pl.: *A szöveg alapján az alábbi kulcsszavakat lehet generálni.*, illetve ***Kulcsszavak:***.

2. Bevezetővel kezdi, majd le is zárja: pl.: *Íme néhány kulcsszó, amely a megadott szöveg alapján generálható: ... Ezek a kulcsszavak segítenek összefoglalni a cikk főbb témáit és koncepcióit.*

3. Csak felsorolja (számozva vagy nem számozva).

4. Valamilyen további szimbólummal, írásjellel számozva sorolja fel (pl.: **).

Továbbá az esetek jórésztében nagybetűvel kezdődnek az általa generált kulcsszavak.

A GPT-4o mini modell által generált kulcsszavak összevetése a szerzői kulcsszavakkal, illetve a PULI Llumix 32K modell által generált kulcsszavakkal az alábbi táblázatban (5. táblázat) látható.

5. táblázat. A teljesen új és a kinyert kulcsszavak további kategorizálása szerinti eredmények

Szerzői kulcsszóval megegyezik (db, %)	PULI Llumix 32K modell által generált kulcsszóval megegyezik (db, %)	Egyikkel sem egyezik (db, %)
129	42	450
20,77%	6,76%	72,46%

A GPT-4o mini modell által generált kulcsszavak esetében annyit vettünk vizsgálat alá, amennyi szerzői kulcsszó is szerepelt a szövegnél. Az összesen 621 db kulcsszóból 72,46% eltért valamilyen módon a szerzői és a PULI Llumix 32K modell által generált kulcsszavaktól. 20,77% megegyezett a szerzői kulcsszavakkal és 6,76% volt, amely csak a másik modell által generált kulcsszavakkal egyezett meg.

A GPT-4o mini modell eredetisége nem csupán kreativitásból fakadt, előfordult, hogy a folyóirat címét adta meg kulcsszóként. Jellemző volt a túl általános kulcsszavak megléte, illetve bizonyos esetekben nehéz volt detektálni, hogy ugyanahhoz a szöveghez adott-e meg kulcsszavakat, vagy csak a gyakori szavakat vette kulcsszónak. A néhány PULI Llumix 32K modellel való egyezés az eltérő írásmódból fakadt, illetve az általánosabb fogalom használatából.

A következőkben néhány példa látható (6. táblázat). Az első lista minden esetben a szerzői kulcsszó, a második a PULI Llumix 32K modell eredménye, míg a harmadik a GPT-4o mini modell által generált kulcsszólista.

6. táblázat. Példák az eltérő és túl általános kulcsszavakra

1.
<i>páros összehasonlítás, előrejelzés, valószínűség számítási háttér, csoportok összefűsülése</i>
<i>páros összehasonlítások, csoportok összefűsülése, előrejelzés, valószínűségi háttér</i>
<i>Thurstone módszer, sporteredmények, női kézilabda, Bajnokok ligája</i>
2.
<i>gamifikáció, oktatás, game engine-ek, szoftverfejlesztés</i>
<i>gamifikáció, oktatás, programozás, játékmotor</i>
<i>Online konferencia, Networkshop 2022, Debreceni Egyetem, HUNGARNET Egyesület</i>
3.
<i>tanárképzési reform, szaktanárképzés, gyakorlóiskolák, képzési kimeneti követelmények</i>
<i>tanárképzési reform, szaktanárképzés, gyakorlóiskolák, képzési kimeneti követelmények, tanárképzés, tanárképző központ, szaktanár, osztatlan képzés</i>
<i>Tanárképzés, Megújulás, Dilemmák, Lehetőségek</i>

Következzen néhány példa, amikor kulcsszóként a terminusok helyett tulajdonnevek jelentek meg.

7. táblázat. Példák az eltérő és túl általános kulcsszavakra

1.
<i>önkéntes katonai szolgálat, felkészítés, kiképzés</i>
<i>haderőfejlesztés, önkéntes tartalékos szolgálat, katonai kiképzés</i>
<i>Önkéntes katonai szolgálat (ÖKSZ), Honvédelmi Minisztérium, Magyar Honvédség</i>
2.
<i>alkotmányosság, értelmiség, jogtörténet, demokrácia</i>
<i>alkotmányosság, értelmiség, jogtörténet, demokrácia</i>
<i>Polgári Szemle, Recenzió, István Kukorelli, Három István</i>

4. Konklúzió

Az elemzés eredményei azt mutatják, hogy a nyelvmodell sikeresen megtanulta, hogy megfelelő számú kulcsszót generáljon az egyes szövegekhez, hiszen átlagosan 4,95 kulcsszót rendel minden szöveghez, míg a szerzők átlagosan 4,67-et. A modell által generált kulcsszavak 40,9%-a és 3,3%-a, tehát összesen 44,2% megegyezik a szerző által megadott kulcsszavakkal. A kulcsszavak 21,8%-a olyan, amely nem található meg közvetlenül a szövegben, ami arra utal, hogy a modell ezek közül sokat saját kreatív módon generált. A teljesen újak között sok a releváns és találó kifejezés.

Bizonyos esetekben a kulcsszavak kategorizálása kihívást jelentett, például a potenciális szinonimák és fordítások esetén, mint pl. *etimológia*–*szótörténet*. Emellett pedig mindenképp nehézséget okoz a szakterületi háttértudás hiánya, mivel a fogalmi struktúrát is annak ismeretében lehet átlátni.

A szerzői kulcsszavakhoz képest az új kulcsszavak típuseloszlása különbséget mutat az első, teljesen új és a harmadik, kinyert kategória között. Az eltérő/nem kategorizálható kulcsszavak aránya mindkét kategóriában közel azonos (54,9% és 57,6%), bár a kinyert kategóriában a kulcsszavak általánosabb fogalmakat reprezentálnak a fogalmi rendszerben, mint például *középkor*, *fejlődés* vagy *pedagógusok*. Az ilyen általánosabb kulcsszavak inkább egy magasabb absztrakciós szintet tükröznek, és ezzel szembe mennek az UNESCO által megfogalmazott specifikusság elvével, ami arra vonatkozik, hogy egy kulcsszónak kifejezetten az adott dokumentumra jellemzőnek kell lennie, és nem lehet alkalmazható általánosan hasonló dokumentumokra. A szerzői motiváció természetesen diktálhat más stratégiát.

A modell által generált teljesen új kulcsszavak specifikusabbak és kreatívabbak, mint azok, amelyeket a nyelvmodell közvetlenül a szövegből nyert ki. Ezek általában többelemű kulcsszavak (vö. összetett terminusok), melyek túlnyomórészt nominális szintagmák és szerkezetek. Ez a tendencia látszik az általánosabb kulcsszó a szerzőjéhez képest kategória arányainál is, mivel 15,1% az arány a teljesen új kulcsszavak 4,1%-ához képest. Ezen túlmenően a teljesen

új kulcsszavak esetében a szerkezetek és a speciálisabb kulcsszó a szerzőjéhez képest kategóriák aránya magasabb (12,3% vs. 2%, illetve 7,4% vs. 5,7%), ami arra utal, hogy a modell generálásából származó kulcsszavak kreatívabbak és speciálisabbak, mint azok, amelyeket kinyert a szövegből a nyelvmódel.

A szerkezeteket azért fontos kiemelni továbbá, mert ilyen jellegű szerkezeteket a szerzői kulcsszavak között is látunk. Például: *emberi és isteni ítélet, katonaegészségügy története, idősoros vs panelregresszió*. Ezek a szerkezetek a tanulmány tematikájának esszenciáját képesek adni, így hatékonyan reprezentálják a szöveget, még ha így nem is tekinthetők terminusnak.

A szerkezetek továbbá a specifikusság elvét követik. A teljesség elve kapcsán, miszerint a kulcsszavaknak le kell fedniük a dokumentum minden olyan témáját, amely potenciális információs értékkel bír, beleértve az implicit, a dokumentumban nem kifejezetten megjelenő témákat is, megjegyzendő, hogy sok tanulmánynál maximalizálva van az adható kulcsszavak száma.

A zero-shot prompting eredményei rámutattak arra, hogy a prompton érdemes lenne módosítani, mivel a kimenet nem egységes, és így közvetlenül nem felhasználható az eredmény. Ezenkívül látható, hogy a GPT-4o mini eredménye bizonytalan, hibázhat. Sokszor nagyon általános kulcsszavakat generál, amelyek nem segítik elő sem a keresőmotorok hatékonyságát, sem a szöveg tartalmának reprezentálását.

A nyelvmódel szerzői kulcsszavakkal történő finomhangolása tehát mindenképpen hasznos, viszont a módellet érdemes további adatokkal finomhangolni annak érdekében, hogy kevésbé operáljon a túl általános, szövegből kinyert kulcsszavakkal.

Az egyelemű kulcsszavak (vö. egyszerű terminusok) kevésbé jól reprezentálják a szöveget (pl. *attitűd, olvasás, értékelés, modellezés*), így érdemes lehet a későbbiekben a többelemű kulcsszavak generálására ösztönözni a nyelvmódellet. Követve a nagy nyelvi modellek fejlődési irányát érdemes lenne az emberi visszajelzésen alapuló megerősítéses tanulási módszertanát (Ouyang et al., 2022) alkalmazni arra, hogy tovább finomítsuk a módellet és arra ösztönözni a nyelvmódellet, hogy absztraktív módon kreatív többelemű, esszenciális kulcsszót generáljon.

A kutatás kétféleképpen is hasznosítható. Egyrészt hozzájárulhat kulcsszógenerálási feladatokra finomhangolt vagy egyéb (például prompttal programozott) nyelvmodellek teszteléséhez, másrészt pedig segíthet a tanítóanyagok további finomhangolásában, hogy a módel a jövőben még jobb eredményeket érhesen el.

Irodalom

- Andonian, A., Anthony, Q., Biderman, S., Black, S., Gali, P., Gao, L., Hallahan, E., Levy-Kramer, J., Leahy, C., Nestler, L., Parker, K., Pieler, M., Phang, J., Purohit, S., Schoelkopf, H., Stander, D., Songz, T., Tigges, C., Thérien, B. & Weinbach, S. (2023). *GPT-NeoX: Large Scale Autoregressive Language Modeling in PyTorch*. [Online]. Available: <https://www.github.com/eleutherai/gpt-neox>
- Barrault L., Biesialska M., Bojar O., Costa-jussà M., Federmann C., Graham Y., Grundkiewicz R., Haddow B., Huck M., Joanis E., Kocmi T., Koehn P., Lo C., Ljubešić N., Monz C., Morishita M., Nagata M., Nakazawa T., Pal S., Post M. & Zampieri M. (2020). Findings of the 2020 Conference on Machine Translation (WMT20). In *Proceedings of the Fifth Conference on Machine Translation. Online: Association for Computational Linguistics*. (1–55).
- Brown T., Mann B., Ryder N., Subbiah M., Kaplan J. D., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D., Wu J., Winter C., Hesse C., Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner C., McCandlish S., Radford A., Sutskever I. & Amodei D. (2020). Language Models are Few-Shot Learners. In *34th Conference on Neural Information Processing Systems (1877–1901)*. Vancouver, Canada: Curran Associates, Inc.
- Devlin J., Chang M.-W., Lee K. & Toutanova K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (4171–4186). Minneapolis, Minnesota: Association for Computational Linguistics. 10.18653/v1/N19-1423
- Dodé Réka (2023). Magyar kulcsszavak vizsgálata és kinyerésük eredményeinek összevetése – pilotkutatás. *Filológia.hu*, 14(1–4), 51–64. <https://doi.org/10.59648/filologia.2023.1-4.3>
- Dodé Réka (2024). Kulcsszavak és terminusok vizsgálata a REAL repozitóriumának anyagán – pilotkutatás. In K. Fogarasi, D. Ittész, & D. Mátyás (Eds.), *Tudásmegosztás, információkezelés, alkalmazhatóság: I. Nyelvhasználat* (58–60). Budapest: Akadémiai Kiadó. https://mersz.hu/dokumentum/m1262tia_60/#m1262tia_58
- Dodé, R. & Yang Z. Gy. (2024a). Further Keyword Generation Experiment in Hungarian with Fine-tuning PULI LluMI 32K Model. In *2024 IEEE 3rd Conference on Information Technology and Data Science (CITDS) Proceedings* (pp. 20-24). Debrecen.
- Dodé Réka & Yang Zijian Győző (2024b). Kulcsszógenerálás magyar nyelvű, hosszú szövegekből nagy nyelvi modellekkel. In G. Berend, G. Gosztolya, & V. Vincze (Eds.), *XX. Magyar Számítógépes Nyelvészeti Konferencia* (257–267). Szeged: Szegedi Tudományegyetem.
- Feldmann, Á., Hajdu, R., Indig, B., Sass, B., Makrai, M., Mittelholcz, I., Halász, D., Yang Z. Gy., Váradi T. (2021). HILBERT, magyar nyelvű BERT-large modell tanítása felhő környezetben. In G. Berend, G. Gosztolya & V. Vincze (Eds.), *XVII. Magyar Számítógépes Nyelvészeti Konferencia: MSZNY 2021* (29–36). Szeged: Szegedi Tudományegyetem, Informatikai Intézet.
- Firoozeh, N., Nazarenko, A., Alizon, F. & Daille, B. (2020). Keyword Extraction: Issues and Methods. *Natural Language Engineering*, 26(3), 259–291.
- ISO. (2022). *ISO 704: 2022 Terminology work — Principles and methods*. <https://www.iso.org/standard/79077.html>
- Jalalov, D. & Gaszicz, K. (2023). *Prompt Engineering Ultimate Guide 2023: Kezdőtől haladóig. Metaverse Post*. <https://mpost.io/hu/prompt-engineering-ultimate-guide/>
- Joulin, A., Grave, E., Bojanowski, P., Mikolov, T. (2017). Bag of tricks for efficient text classification. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. 427–431. Valencia, Spain: Association for Computational Linguistics.
- Jurafsky, D. & Martin, J. H. (2023). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (3rd ed.). (Letöltés: 2024. október 10.) https://web.stanford.edu/~jurafsky/slp3/ed3book_jan72023.pdf
- Liu, R., Lin, Z., & Wang, W. (2021). Addressing extraction and generation separately: Keyphrase prediction with pre-trained language models. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29, 3180–3191.

- Mikolov, T., Chen, K., Corrado, G. & Dean, J. (2013). *Efficient estimation of word representations in vector space*. arXiv. <https://doi.org/10.48550/arXiv.1301.3781>
- Nemeskey, D. (2020). Egy emBERT próbáló feladat. In *XVI. Magyar Számítógépes Nyelvészeti Konferencia: MSZNY 2020* (409–418). Szeged: Szegedi Tudományegyetem, Informatikai Intézet.
- Nemeskey, D. (2021). Introducing huBERT. In *XVII. Magyar Számítógépes Nyelvészeti Konferencia: MSZNY 2020* (3–14). Szeged: Szegedi Tudományegyetem, Informatikai Intézet.
- Nomoto, T. (2023). Keyword extraction: A modern perspective. *SN Computer Science*, 4 (92), 1–15. <https://doi.org/10.1007/s42979-022-01481-7>
- Ouyang L., Wu J., Jiang X., Almeida D., Wainwright C. L., Mishkin P., Zhang C., Agarwal S., Slama K., Ray A., Schulman J., Hilton J., Kelton F., Miller L., Simens M., Askell A., Welinder P., Christiano P., Leike J. & Lowe R. (2022). Training language models to follow instructions with human feedback. In *Proceedings of the 36th Conference on Neural Information Processing Systems*. New Orleans, Louisiana.
- Reynolds, L. & McDonell, K. (2021). *Prompt programming for large language models: Beyond the few-shot paradigm*. arXiv. <https://doi.org/10.48550/arXiv.2102.07350>
- Sammet, J. & Krestel, R. (2023, September). Domain-Specific Keyword Extraction using BERT. In *Proceedings of the 4th Conference on Language, Data and Knowledge* (659–665).
- Sun, Y., Qiu, H., Zheng, Y., Wang, Z. & Zhang, C. (2020). SIFRank: a new baseline for unsupervised keyphrase extraction based on pre-trained language model. *IEEE Access*, 8, 10896–10906.
- Syed, M. & Gadesha, V. (2025). *What is zero-shot prompting?* IBM, 29 January 2025. Megtekintve: 2025. március 30. <https://www.ibm.com/think/topics/zero-shot-prompting>
- Szakadát István, Szóts Miklós & Szaszko Sándor (2006). *Az ontológia fogalma, építése, kezelése*. Megtekintve: 2024. október 10. <http://www.ontologia.hu/ontologies.pdf>
- Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., Bikel, D., Blecher, L., Chen, M., Cucurull, G., Esiobu, D., Fernandes, J., Fu, J., Fu, W., Fuller, B. & Scialom, T. (2023). *Llama 2: Open Foundation and Fine-Tuned Chat Models*.
- UNESCO. (1975). *UNISIST Indexing Principle* (SC. 75/WS/58).
- Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł. & Polosukhin I. (2017). Attention Is All You Need. In *Proceedings of the 31th International Conference on Neural Information Processing Systems*. Long Beach, CA, USA: Curran Associates Inc. 6000–6010.
- Yang, Z. Gy., Novák, A., Laki, L. J. (2020). Automatic Tag Recommendation for News Articles. In *Proceedings of the 11th International Conference on Applied Informatics (ICAI 2020)* (442–451). Eger.
- Yang, Z. Gy., Dodé, R., Ferenczi, G., Héja, E., Jelencsik-Mátyus, K., Körös, Á., Laki, L. János, Ligeti-Nagy, N., Vadász, N., Váradi, T. (2023a). Jönnék a nagyok! BERT-Large, GPT-2 és GPT-3 nyelvmodellek magyar nyelvre. In *XIX. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2023)* (247–262). Szeged: Szegedi Tudományegyetem, Informatikai Intézet.
- Yang, Z. Gy., Laki, L. J., Váradi, T. & Prószéky, G. (2023b). Mono- and multilingual GPT-3 models for Hungarian. In *Text, Speech, and Dialogue*, Lecture Notes in Computer Science (94–104). Plzeň, Czech Republic: Springer Nature Switzerland.
- Yang, Z. Gy., Dodé, R., Ferenczi, G., Hatvani, P., Héja, E., Madarász, G., Ligeti-Nagy, N., Sárossy, B., Szaniszló, Zs., Váradi, T., Verebélyi, T. & Prószéky, G. (2024). The First Instruct-Following Large Language Models for Hungarian. In *2024 IEEE 3rd Conference on Information Technology and Data Science (CITDS) Proceedings* (247–252). Debrecen: University of Debrecen.

ADRIANNA KOSTIC

University of Szeged
Szeged, Hungary
<https://orcid.org/0009-0007-4032-3997>

Adrianna Kestic: Minorities in the digital space: Translanguaging practices of Vojvodina Hungarians online

Alkalmazott Nyelvtudomány, XXV. évfolyam, 2025/1. szám, 84–103.

[doi:http://dx.doi.org/10.18460/ANY.2025.1.005](https://doi.org/10.18460/ANY.2025.1.005)

Minorities in the digital space: Translanguaging practices of Vojvodina Hungarians online

This small-scale study aimed to gain insight into Vojvodina Hungarians' digital habits and digital language practices from a qualitative perspective as an extension to a previous study on the same topic (Kestic, 2024). Approached from a slightly different perspective, semi-structured interviews were chosen as the method of data collection, conducted with some of the same participants as those in Kestic (2024). The Vojvodina Hungarian participants in the present study displayed differing life stories in relation to technology (cf. Lee, 2014), but with several noteworthy overlaps regarding their experience with the digital era, their language choices and factors affecting those, as well as their translanguaging. The results show that the participants are highly digitally oriented, often turning to the internet first to seek information and entertainment, to communicate and connect with others, or run errands digitally. Multilingual practices are common among the participants and are integral parts of their speech and written texts. Their language choices are often dependent on environmental, social, and frequently personal factors (including personally perceived skills and knowledge, intention of messages, membership, belonging, and identity), especially when it comes to translanguaging as an identity marker.

Keywords: digital language use, translanguaging, language choices, multilingualism, Hungarian minority

1. Introduction

Increasingly, scholarly interests are shifting towards analyzing and making sense of language use in combination with social life and identity, with the term 'translanguaging' attracting primary focus. Multilingualism and translanguaging support some of the same ideas, whereby the speaker transcends linguistic boundaries and uses multiple languages in communication regardless of their proficiency in those (Aleksić and García, 2022; Androutsopoulos, 2015). Multilingualism aims to erase the boundaries between languages that have been socio-politically constructed and separated, and, instead, views languages as interconnected complex systems (Androutsopoulos, 2015; Lee, 2016; Li, 2011). In this sense, translanguaging is "the deployment of a speaker's full linguistic repertoire without regard for watchful adherence to the socially and politically defined boundaries of named (and usually national and state) languages" (Otheguy et al., 2015, p. 283). Translanguaging also has a social dimension, where the choice to translanguage is often intentional and serves as an identity marker

(Androutsopoulos, 2015; Li, 2011). Due to monolingual ideologies that view languages in isolation and reject non-normative language use, multilingual communication has long been stigmatized, despite being an integral part of people's communicative practices worldwide (Aleksić and García, 2022). However, that is not to say that there have not been instances where multilingual practices have been recognised and initiatives made in favour of them. For one, the case of Yugoslavia's language education program, between the 1960s and 1980s, is a great example that Mandić and Rácz (2023) look into, which strived to support multilingualism considering Yugoslavia's multiethnic composition (Mandić and Rácz, 2023, p. 461) by introducing the LSE (language of the social environment, e.g. Hungarian in majority communities like Serbian in minority communities) as an optional second language class in schools in areas of Yugoslavia that had a diverse ethnolinguistic population. One of the most multilingual areas of Serbia is Vojvodina (Ferdinand and Komlosi, 2017; Trombitás and Szügyi, 2019), where one of the largest cities, Subotica, is home to as many as fourteen languages spoken as native tongues, many with an official status (Ferdinand and Komlosi, 2017; Statistical Office of the Republic of Serbia, 2012; Trombitás and Szügyi, 2019). The aim behind introducing the LSE classes was to promote bilingualism among monolingual majority individuals, who would ideally also become bilinguals like most minorities in the territory of Yugoslavia. Mandić and Rácz (2023) have found that the ideology behind LSE views languages as resources for societal multilingualism and, at the same time, as a solid basis for national unity, and discussions about the reinstating of the subject began in the 2010s signaling just how vital educational, societal, and institutional support and community are.

The present small-scale study aims to gain insight into the online linguistic practices and translanguaging of Vojvodina Hungarians, focusing on a topic similar to that of Kostic (2024), which approached the same research questions using different data collection tools and methods. To approach some of these questions from a slightly different perspective, semi-structured interviews were conducted with some of the same participants who provided data for the previous study (Kostic, 2024). The present study aims to answer the following questions:

1. For what purpose and how do Vojvodina Hungarians use their language in digital spaces?
2. What factors influence the language use of Vojvodina Hungarians in digital spaces, particularly in terms of language choices and translanguaging?

Section 2 below provides an overview of previous literature on topics such as the intricacies of online language practices and the potential factors influencing them, the techno-biographical interview, as well as earlier studies on the digital presence of minority Hungarians, the importance of identity, and digital language habits.

2. Literature Review

2.1. Language, participation, and identity in the digital age

Digital language use is influenced by numerous factors. A very decisive factor that Kornai (2013, p. 1) describes is a language's digital vitality, "whereby a language increasingly acquires digital functions and prestige as its speakers increasingly acquire digital skills," which does not happen at the same rate for everyone. Additional factors such as age, occupation, and personal interests are also important. At the same time, active online participation, language choice, online communities, and websites one chooses to visit and be a part of are also closely related to identity. Two studies by Lee (2014) and Kelly-Holmes (2004) illustrate just how deeply interwoven these aspects are with one another.

To begin with, in a 2014 study, through techno-biographical interviews and recordings of social media browsing done by her participants, Lee (2014) examined the situated online language practices and online identities of Cantonese and English bilingual undergraduates in Hong Kong. By focusing on habits related to internet and daily technology use, she uncovered how her participants utilized many online resources to convey personalized messages, posts, and content, which all played a crucial part in identity play. As Lee puts it, techno-biography is a "life story in relation to technologies" (2014, p. 94), where participants share their own stories, as they experience technology both in the present and the past (Ching and Vigdor, 2005; Lee, 2014). Lee (2014) concludes that, although there ought to be overlaps, people's experience with technology is still very individual, which is also true for linguistic habits that heavily depend on not only context but also the audience one wishes to address and communicate with.

Apart from status updating, commenting, post creation, sharing photos, and creating a profile online (Kelly-Holmes, 2004; Lee, 2014), the availability of all sorts of linguistic and technological resources (keyboards, different input languages, creative spelling, mixing styles and spoken languages, special characters and symbols from other languages one might not even speak) are key elements when it comes to building and expressing one's identity online (Androutsopoulos, 2007; Lee, 2014; Tagg and Sargeant, 2012; Vaisman, 2011). Due to the situatedness of identity, studying the identity practices and lived experiences of internet users online requires a context-conscious approach (Lee, 2014).

Another study, by Kelly-Holmes (2004), looked at students' internet practices from multiple countries (Italy, Macedonia, Indonesia, Poland, France, and Tanzania), and discovered that the choice to use English over one's first language on the internet was not as prevalent as one would think, considering the overwhelming English content present in the online space. The few cases in which English was chosen over the L1 included situations where the respondents were from countries (for example, Macedonia, which had recently gained

independence at the time) that lacked the financial means to establish and maintain websites in their own respective majority languages. In combination with language choice, identity is another factor to consider when it comes to digital media consumption and content creation. Identity is understood to be fluid, "like masks that can be worn and taken off in different contexts of social interaction" (Lee, 2014, p. 91, commenting on Goffman, 1990[1959]). This understanding applies whether the interaction occurs online or offline. Very often, these "masks" are worn intentionally with specific aims behind them, such as emphasizing inclusivity, showing support, or even strengthening one's community online when it might not be possible to do so in person (Lee, 2014; Molyneaux et al., 2014; O'Carroll, 2013). What we can establish from the studies above is that with the help of various textual, linguistic, typographical, and visual resources and tools, users can creatively convey personalized messages and shape their own identities as they wish in the digital space.

2.2. Translanguaging and multilingualism in minority communities

As opposed to code-switching, translanguaging and multilingualism understand languages as interconnected systems instead of separate linguistic systems (named languages) with clear boundaries (García and Otheguy, 2019; Otheguy et al., 2015). Very often, Second Language Acquisition studies as well as classroom environments foreground dominant language varieties, whereby an idealized native speaker model is promoted (Cenoz and Gorter, 2017, 2019), due to the monolingual bias, a socio-political byproduct. An increasing number of scholars have started criticizing these traditional views and approaches, and are shifting their focus on exploring multilingual communities, where instead of looking at people's competence in languages, they are instead making sense of speakers' 'multilingual repertoires' (Cenoz and Gorter 2017, 2019) as a whole interconnected system. These more contemporary perspectives aim to provide a nuanced understanding of a person's spoken languages, encompassing both linguistic and semiotic resources, regardless of the individual's level of proficiency in those languages. This approach differs from viewing languages in isolation or identifying weaknesses in one's repertoire. For these reasons, translanguaging as a conscious practice is promoted, as it not only allows the speaker to fully utilize all of their linguistic, pragmatic, and semiotic language knowledge, but also gives them an opportunity to cultivate and portray their desired identities, which is especially prevalent in minority settings (Cenoz and Gorter, 2017, 2019). While translanguaging promotes and fosters the use of minority languages and in a way challenges language hierarchies by viewing a person's spoken languages as equals regardless of their hierarchical position in a particular context, scholars have also been raising concerns about the potential adverse effects that the promotion of translanguaging could bring about if social contexts and power dynamics between the languages and the speakers are not

taken into careful consideration. For minority languages to survive and thrive, there is a need for supportive communities and places where the use of the minority language becomes necessary. This is especially important to establish in pedagogical institutions (Cenoz and Gorter, 2017, 2019; García and Otheguy, 2019; Otheguy et al., 2015), as they play a crucial role in developing speakers' awareness of the necessity for their minority languages (Cenoz and Gorter, 2017, 2019). Many studies have shown that minorities can use the internet to their advantage and further strengthen their ties with their community, but also enrich and even strengthen their identities and language use in creative and novel ways, whether it is by translanguaging or mixing languages (Jongbloed-Faber et al., 2016; Lee, 2014; Lynn et al., 2015). While the latter two have received criticism for resulting in weakened language varieties, the experiences shared by individuals from these groups are encouraging and show the desire to actively participate in online content creation, activities, and communication (Pauwels, 2016; Tagg, 2015). Belmar and Glass (2019, p. 5) argue that the once traditional language communities have been deeply affected by communication technology, as it "forces us to re-contextualize languages, especially [...] minority languages," due to the fact that the internet provides a space for everyone regardless of one's mother tongue, geographical location, or social position.

Studies that were particularly conducted in Serbia have mainly focused on topics like the impact of digital media (Đerčan et al., 2020; Pásztor-Kicsi, 2012, 2016), digital media consumption (Pásztor-Kicsi, 2016), attitudes towards different varieties of languages in the region (Kovacs Rac and Halupka-Rešetar, 2018; Kovács Rác, 2012), and the multilingual situation and environment (Balla et al., 2012; Pilipenko, 2016).

One study focused on the impact of English language media and the internet on the language use and identity of Vojvodina Hungarians (Pásztor-Kicsi, 2016). According to Pásztor-Kicsi (2016), digital media did not appear to have a significant impact on the respondents' attitudes towards their spoken varieties or English. Regarding the Vojvodina Hungarians' attitudes towards different varieties of Hungarian in the region, including Hungary (Jánk and Rási, 2023; Kovacs Rac and Halupka-Rešetar, 2018; Kovács Rác, 2012), the collected data reveal similar trends. While it was noted that participants occasionally experience negative attitudes or discrimination against regional varieties (Jánk and Rási, 2023; Pásztor-Kicsi, 2016), the vast majority of respondents reported having neutral or positive attitudes towards their own spoken variety and the standard Hungary Hungarian. The above studies inquired about language attitudes in younger individuals, such as 5th- to 8th-grade students (Kovacs Rac and Halupka-Rešetar, 2018) and university students (Pásztor-Kicsi, 2016), as well as older individuals (Jánk and Rási, 2023).

In a 2011 research project, *Mozaik2011*, Szabó et al. (2013) found drastically different stances towards the importance of national belonging, identity, and the

value of one's mother tongue among minority Hungarians living outside of Hungary. As opposed to the Hungarian minority in Austria, the Hungarian minority in Vojvodina (who often face financial difficulties and linguistic and social discrimination) find their identity and mother tongue to be of utmost importance in community building and cultural, ethnic, and linguistic preservation, especially in online settings. Within the same set of studies (*Mozaik2011* in Szabó et al., 2013), Galács and Ságvári (2013) examined how the internet influenced the lives, language use, international and intranational connections, and identities of minority Hungarian youth living outside Hungary in the neighboring regions. Their results revealed that already at the beginning of the 2010s, young people were undoubtedly using the internet and social media and often expressed positive attitudes towards multilingualism for the expanding possibilities it can bring about. The participants also discussed the importance of knowing their country's dominant language, particularly for educational and economic reasons. Interestingly, a minority of young Hungarian people expressed a variety of opinions regarding the practice of language mixing online, as some found it distressing and confusing when they did not understand the language and expressions being used. Their respondents had extensive experience with digital media, and global, international, and primarily English influences were also observed in terms of language choices and content consumption (Galács and Ságvári, 2013). With somewhat similar findings to Galács and Ságvári's (2013), in a recent paper on the digital and linguistic practices of Vojvodina Hungarians (Kostic, 2024), the results uncovered that multilingual language practices are an integral part of their daily lives (see also Mandić and Rác, 2023), both online as well as offline, however, Hungarian and English are in most cases and contexts preferred over Serbian most probably due to the participants' low confidence in their Serbian knowledge and their own perceived language skills.

Research to date on Vojvodina Hungarians thus illustrates how the use of digital platforms, experience with digital media, language attitudes, and identity all affect and constantly shape one another. Although Vojvodina Hungarians continue to face challenges such as linguistic and social discrimination, they nevertheless express their desire to prioritize the preservation of their language and culture, which is especially taking place more and more often in online settings. These findings also highlight the importance of considering factors such as language choice, digital literacy, and cultural identity in understanding the experiences of minority language communities in the digital age.

3. Methodology

3.1. Data collection and analysis

The present small-scale study invited Vojvodina Hungarians of all ages, genders, and occupations to share their stories and experiences through semi-structured interviews. The participants who took part in the study were 12 Vojvodina Hungarians from the same group of participants who filled in the questionnaire reported by Kostic (2024). The interview data were analyzed individually, and emergent themes were established. The questions and sections of the questionnaire (Kostic, 2024) were adapted to the interview format, where the questions served solely as a guide during the 30-to 40-minute interview sessions. The participant selection process was random, as the questionnaire from Kostic (2024) was shared on Facebook and other social media platforms, including Instagram and Reddit. Anyone could complete the questionnaire and indicate at the end whether they wished to participate in an additional interview with similar discussion topics. Those who decided to give interviews were asked to leave a contact email or a link to their Facebook or Instagram page, where they were later contacted to arrange an online interview session via Facebook Messenger or an Instagram call. The only criterion was that the participants were from Vojvodina and spoke Hungarian as their mother tongue. Considering that the questionnaire was shared on multiple platforms, including various Facebook groups, it was not unlikely that the participants would not know the researcher personally, but in the end, none of the 12 participants were familiar with them. They also did not have any contact with the other participants, as each interview was done separately and at different times. The interviews were recorded with the participants' prior consent for easier analysis later on, but were subsequently deleted once the recordings were transcribed.

Like the questionnaire itself in Kostic (2024), the interview was divided into several sections, each covering a different discussion topic. All the interview questions used during the interviews are listed in the Appendix below. The interviews were conducted using the same set of questions. They were semi-structured in nature, providing participants with more freedom and opportunity to express their attitudinal standpoints, share their digital and linguistic habits (Appendix, sections *B* and *C*), and their overall experience with language use and social media in the Vojvodina Hungarian setting (Appendix, section *D*). The second section (Appendix, Section *B*) in particular covered their history with digital devices, the internet, and social media by incorporating Lee's (2014) techno-biographical approach.

3.2. Participants

The 12 participants who gave interviews represented diverse backgrounds, ages, and had varying digital habits. All interviewees were given random names for easier identification within the present study and to also ensure their anonymity. In an earlier study (Kostic, 2024), six age groups were established (under 18, 18–25, 26–35, 36–45, 46–55, and over 55) to more easily distinguish any potential differences regarding digital habits, shared experiences, and linguistic habits. Since there were no participants under the age of 18, the other five age groups will be considered from this point on. The first section of the interview (Appendix, *A. Background information*) uncovered that there were 4 participants within the age categories under 18 and in the 18–25 range, another four individuals in the 26–35 and 36–45 ranges, and four more individuals within the 46–55 and over 55 age groups.

The youngest participants included Dávid, Dorina, Dorottya, and Dániel, all of whom were between the ages of 18 and 25. They were all had similar digital habits in terms of active online participation, all attended university and were in the following fields: Dávid was studying education at the University of Szeged, Dorina was studying art also at the University of Szeged, while Dorottya was studying IT at the same university, and lastly, Dániel was studying biology at the University of Novi Sad. Another four participants (Emese, Erika, Éva, and Endre) who gave interviews belong to the second two age groups (26–35 and 36–45). They were working full-time in Vojvodina, Serbia, with Emese being in her third year of IT at the University of Novi Sad. Erika had a high school diploma and was working as a sales assistant at the time of the interview in Horgos, Éva had a bachelor's degree in business and administration, and was working at an administrative office in Subotica. In contrast, Endre held a master's degree in psychology and worked as a psychologist in Novi Sad. The last four participants (age groups: 46–55 and over 55) were Luca who had a high school diploma and was working as a nail technician in Novi Sad, Levente, who finished vocational school and was a car mechanic in Kikinda, Liza, who attained her master's degree in education at the University of Novi Sad and was a high school teacher in Zrenjanin, and Leon, who had a PhD and was a retired private investigator in Subotica.

All of the interviewees are ethnically Hungarian, speak Hungarian as an L1, and some also speak Serbian well (Leon, Liza, Endre, Erika, Éva, Emese, Endre, and Dániel), and six of them (Dániel, Dávid, Dorottya, Dorina, Emese, and Endre) speak English as an additional language they learned either at school or through the internet. They all share the same mother tongue (Hungarian) and own at least two digital devices, mainly smartphones, smart TVs, and a PC or a laptop. They also shared that they have constant internet access and spend up to 5-6 hours on the internet on a daily basis, during which they consume all sorts of social media (mainly Facebook, TikTok, and Instagram). They tend to communicate (mostly

using Viber and Facebook Messenger) and look up information on the internet much more frequently nowadays as opposed to 10-15 years ago, when their primary sources of information were television, radio, newspaper, or seeking information directly from other people.

4. Results and Discussion

The present section is divided into three subsections, each addressing a distinct research question, and the results and observations from the interviews have been categorized accordingly.

4.1. The Vojvodina Hungarians' digital language use

Two sections in the interview (see Appendix, sections *B. Linguistic habits, attitudes, and spoken languages*, and *D. Online communities in Vojvodina*) focused on gathering detailed responses, anecdotes, and examples to answer the first research question, which focuses on the linguistic habits of the Vojvodina Hungarians in digital spaces. As mentioned above in Section 3.2, most participants speak Serbian and English well, having learned these languages either at home from their families and friends, or in school. They mentioned additionally that they often switched between Serbian and Hungarian (younger individuals also often mentioned English too) in their day-to-day conversations depending on the type of the situation, the intended message they wished to convey, and often the power relations among interlocutors, which tended to also extend to online communicative settings especially in private messages taking place on Instagram, Viber, WhatsApp, and Facebook Messenger mainly.

The 12 participants regularly use Hungarian on the internet, very often in combination with Serbian and English, while the younger ones also tend to combine all three languages when communicating. When asked about their knowledge, their perceived skills, and attitudes towards the three languages, the participants generally tended to say positive things regarding the languages, however, a few individuals mentioned that they were not too comfortable speaking Serbian (Dorina, Dorottya, Dávid) for more extended periods nor reading Serbian Cyrillic, or English (Leon, Liza, Endre, Erika, Éva, Emese, and Dániel) also due to their low proficiency in the language or their lack of knowledge of the language. Several participants expressed that they were proud of being a Vojvodina Hungarian and many were also fond of the Vojvodina Hungarian variety due to the "colourfulness of the language" (Éva) and also for the "uniqueness of its vocabulary" (Levente), which many agree is more creative and likeable than most Hungary Hungarian varieties, especially the standard (Leon, Dorina, Dániel, and Éva). Almost all participants mentioned that they frequently use Serbian words and expressions in their Hungarian speech, which they attribute to a lack of knowledge of adequate Hungarian words. Éva and Leon also stated that it is much easier to speak this way than to try to find the adequate

Hungarian words, as it would only cause more misunderstandings during conversations, considering that most people are already accustomed to Serbian, English, and even Croatian terms.

Section *D* (see Appendix) of the interviews also revealed that participants regularly visit Vojvodina Hungarian Facebook groups to share or read about fellow Vojvodina Hungarians' experiences and to connect with people who share similar views. Their experiences in these groups vary between positive and neutral, however, both Dániel and Emese mentioned that they often saw posts and comment sections in some larger Hungarian Facebook groups (group orientations frequently include cooking and baking, marketplaces, and travel) where minority Hungarians, not only from Vojvodina but all around Hungary's outside and inside borders, are corrected and ridiculed for their language use, choice of vocabulary, and even sometimes for their origins usually by majority Hungarians from Hungary. Dániel also mentioned that these acts of discrimination are occasionally present in Vojvodina Hungarian Facebook groups, but to him, they are seemingly much less frequent (similar findings in: Jánk and Rási, 2023; Pásztor-Kicsi, 2016).

Lastly, to address the second research question regarding the factors that affect their language choices in digital settings, the participants generally mentioned the same things as Kostic (2024). The third section (Appendix, section *C. Digital habits and language use online*) in the interviews that focused on digital habits and general language use tendencies online unveiled that, generally, those who spoke languages other than Hungarian well (i.e. mainly English and Serbian), consumed content and communicated in those languages on a regular basis. Most frequently, their own language choices depend on the language they saw in comments or online content, but the type or topic of the content was also often the most decisive factor. The 12 participants also mentioned that there are situations, where they intentionally switch between languages or choose another language than that of the comments, posts, or content, to express their linguistic, national, and cultural identities (see also Androutsopoulos, 2015; Li, 2011) and simultaneously indicate their ties to Vojvodina Hungarians, Serbia, or even the Balkan. Similarly to online settings, translanguaging and switching between the languages one speaks are also integral to the face-to-face conversations of Vojvodina Hungarians (cf. Aleksić and García, 2022; Mandić and Rácz, 2023). As Emese states, it very often happens that conversations take place in two or sometimes even three languages at once:

"Here in Vojvodina, it's quite an everyday thing that sometimes we speak Hungarian, sometimes we speak Serbian, but there are occasions where we continuously use both of the languages, and it really depends on what way of speaking is easier. The other day I was at the market and started speaking in Serbian out of habit. Considering that there are more Serbs, we sometimes tend to carry out whole conversation in Serbian even with people we don't know are Hungarians, but that's okay, this is what we are

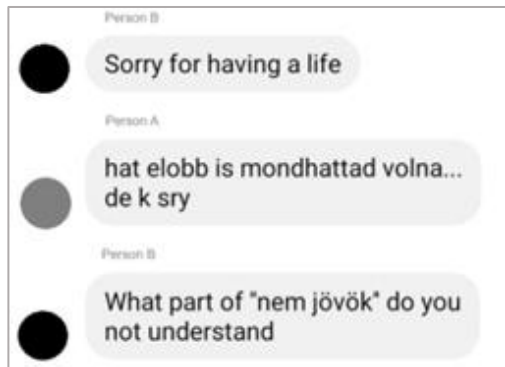
used to it and there's nothing wrong with that. But there have also been times when I spoke Serbian with someone I knew was a Serb, and they answered me back in Hungarian, and the whole conversation went on like that." (Quote 1, Emese)

As she explains, this can happen out of habit (starting conversations in Serbian due to it being the majority language), or because it is easier for interlocutors to communicate in such multilingual ways. However, it also very frequently happens due to people's desire to show respect for the other person's mother tongue. Being able to communicate digitally facilitates the maintenance of familial, work-related, and other connections, surpassing physical obstacles. Another critical function that digital spaces serve for the Vojvodina Hungarians is providing access to a wide range of information, from the local and more personal to the global level.

4.2. The Vojvodina Hungarian way of translanguaging

The participants also often found themselves not only switching between languages in conversations, both online and offline, but also intentionally translanguaging to signal their belonging to specific communities (cf. Cenoz and Gorter, 2017, 2019). In the following, three figures contain translanguaging examples that Dorina and Dániel shared with me at the end of the interview. Figures 1, 2, and 4 below are examples of conversations that took place on Facebook Messenger, while Figure 3 is a public post Dorina posted in the *Magyarországon tanuló vajdasági diákok* "Vojvodina Hungarian students studying in Hungary" Facebook group. All of the figures display the screenshots in their original form, along with English translations next to them and * indicating the language each message was written in (* for English, ** for Hungarian, and *** for Serbian). The Facebook post (Figure 3), as well as most of the messages, were written in Hungarian, with the exception of the first and last messages in Figure 1, which were written in a combination of English and Hungarian. Figure 4's messages are in Hungarian and Serbian. In contrast, the conversation in Figure 2 took place in all three languages (Hungarian, Serbian, and English). Two Messenger conversations were on the topic of planning an in-person meetup among friends in Figure 1, and among old classmates from high school in Figure 2, the Facebook post was written in relation to education and degree accreditation, and Figure 4's series of messages were responses discussing what it is like to be a Vojvodina Hungarian and being able to understand inside jokes, and what it means to the commenters to share a common language, culture, and lived experience with fellow Vojvodina Hungarians.

To begin with, the screenshot Dorina shared with me (Figure 1) was a conversation between her and her university friend, who also spoke English well.

Figure 1. Dorina's (Person B) private messages on Facebook Messenger with her university friend.

Translation: (underlined parts are written in Hungarian in the original)

Person B: *Sorry for having a life*

Person A: *Well, you could have told me earlier, but ok sorry [for saying anything]*

Person B: *What part of I am not coming do you not understand*

Dorina also emphasized that they usually converse in this manner in person, too, as it is much more entertaining and often easier to express themselves through translanguaging rather than speaking only one language. Additionally, Dorina and Dániel also mentioned that their language choices and translanguaging often depend on the topic of their conversations, but it usually comes naturally (see also Aleksić and García, 2022), and they do not pay too much attention to it.

Like Dorina's example, in Figure 2 below, we can see that Dániel's too includes instances of translanguaging from multiple people involved in the Messenger group conversation, where words like the Serbian *vikendica* "holiday home" or the English expression 'sooner or later' is used. Notably, quite a few Serbian words (and some Hungarianized Serbian words) were also uttered during the interviews, such as the Hungarianized *centárba* "to the [city] center" from Serbian *centar* "city center," and the Serbian *firma* "firm."

Figure 2. Dániel's (Person B) private messages on Facebook Messenger with a group of old high school peers planning a reunion.

Translation: (underlined parts are written in Serbian in the original)

Person A: *But who has a holiday home? 🤔*

Person B: *We will find out soon 😊*

Person C: *Sooner or later*

We could rent Rage (bar/pub) in Ada

Person B: *Is that a holiday home?*

Person C: *It's where my birthday was at*

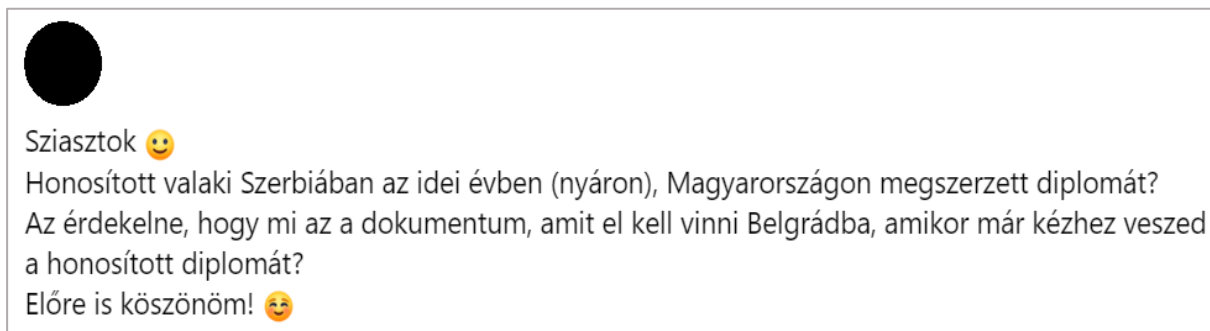
Person D: *Where can you BBQ there?*

On the other hand, Dorina mentioned that there are certain situations and topics where she has to pay attention to her writing and choice of words:

"I think it is much more interesting and special if a person knows how to mix multiple languages into their speech [instead of just speaking in one language]. Personally, I often tend to combine Serbian words into my speech intentionally, especially when I'm in an environment where I want to bring attention to myself or my origins, and honestly, it's the same with English too. On the internet, specifically on TikTok, I often mix English with Hungarian, but also Serbian because since TikTok became famous, for some reason the Balkans have become very cool and we have to enjoy and take advantage of this position/privilege. But of course, it's not always like that, if it's a serious topic I'm much more careful with what I write and in what language I am writing." **(Quote 2, Dorina)**

Although Messenger messages are also written language, they are typically rapid responses that resemble spoken language, which makes them distinct from public Facebook posts, such as Dorina's in Figure 3.

Figure 3. Dorina's public Facebook post in the Magyarországon tanuló vajdasági diákok "Vojvodina Hungarian students studying in Hungary" Facebook group.



Translation: *Hello everyone! 😊 Has anyone accredited their degree that they obtained in Hungary this summer in Serbia? I am wondering which document I need to take with me to Belgrade when picking up my accredited degree. Thank you in advance! 😊*

If we take a closer look at the text in Figure 3, we can see that it is carefully constructed and written in a much more formal way than her messages in Figure 1 above. However, it is also noteworthy that this is an entirely monolingual post, where she highlighted that she also paid special attention to write the Hungarian name of Belgrade, *Belgrád*, as opposed to the Serbian *Beograd*, which she usually tends to use when referring to the capital city, even when speaking and writing Hungarian.

Lastly, Figure 4 below is a screenshot Dániel provided, which was a series of messages reacting to a message from another Vojvodina Hungarian within a larger Vojvodina Hungarian Messenger group chat. The message they were reacting to

was a humorous short text written in Hungarian by Dániel (the first message in blue in Figure 4), filled with Serbian borrowed words and expressions that elicited mixed responses. Despite the original message being written in Hungarian, the people responding, including Dániel (Person A in Figure 4), sometimes chose to reply in Serbian and instead combine Hungarian words and expressions into their comments. Dániel clarified that his linguistic choices were deliberate when writing his responses and were intended to signal that those who spoke both Hungarian and Serbian could easily understand one another, even if the messages contained translanguaging.

Similar reasons were mentioned by Éva and Leon, who talked about communication with other Vojvodina Hungarians in face-to-face situations. Dániel's translanguaging instances included the use of the Serbian word *ljudi* "people," the Hungarianized version of the Serbian *gužva* "crowded," the Hungarianized version of *granica* "border," the Hungarianized version of *svercer* "smuggler of goods," the Hungarianized version of the Serbian *buvljak* "flea market," and the Hungarianized version of the Serbian *vikend* "weekend." In this conversation on Facebook Messenger, four out of five individuals expressed that they regularly engage in translanguaging, while Person E and, to some extent, Person B expressed their disagreement with this way of speaking. As we can see, Person D and later Dániel (Person A) too decided to mix in Serbian when responding to Person E's and Person B's messages, which Dániel was meant to signal that he does not necessarily differentiate between his two spoken languages here, and instead, views both as equal and equally important.

Figure 4. A series of messages reacting to a message from another Vojvodina Hungarian within a larger Vojvodina Hungarian Messenger group chat shared by Dániel (Person A).



Translation: (underlined parts are written in Serbian in the original)

Person A: *People, there is such a crowd here at the border that I have never seen so many [goods] smuggles before... the Hungarians will have a reason to come to the flea market this weekend* 🎨👍

Person B: *Oh, what nonsense*

Person C: *What's wrong with it?! We also speak like that at home, only the other way around!!! We weave Hungarian words into our Croatian mother tongue. I am proud of this!!!!*

Person D: *Now that's real Hungarian speech oof, I'm going to go the flea market then tomorrow* 👍 *thanks for the information*

Person E: *what are you talking about here, this doesn't look like anything*

Person A: *So what? It's not like you didn't understand what I wrote | Come on, man*

Person D: *Please, don't act like you don't actually speak like that in reality | Well f*ck, that's how it is | The more languages you speak, the more you're worth – that's what my grandpa used to say. | I wish you a nice day! (original: German)*

As mentioned above, these Figures include messages mainly in Hungarian with occasional English and Serbian words that have been intentionally integrated into the messages. Aside from Dániel (see quote four below) and Dorina (see quote two above), Dorottya also shared a few reasons why she intentionally chooses to engage in translanguaging, in these messages, particularly, as well as in general, online or offline:

"We constantly mix together multiple languages when talking with my Vojvodina Hungarian girl friends. The funniest thing is that we can easily understand one another even if our messages are a mix of Hungarian, English, and Serbian words in the group chat or when we talk on Facetime."
(Quote 3, Dorottya)

While Dániel did mention that he uses words and expressions from Serbian and English (and rarely from German too) in cases where he cannot recall the word or phrase he is looking for, there are also various situations where he intentionally switches to English and/or Serbian when he intends to get across a message with underlying meaning often to signal particular aspects of his identity, whether that is his national, linguistic, or shared cultural identity:

"It often happens that I can't recall a word or a term [in Hungarian], and that is the reason I say it in Serbian, sometimes also in English, but that really depends on who I'm talking to. If I know that the person doesn't know or speak English or Serbian, I try not to throw in words into my speech in that language, but sometimes I can't help it and they slip out. However, I also have specific moments when I do it fully consciously. This actually happened at a football match in Hungary, when the Serbs were playing against the Hungarians and I was cheering for Serbia, because I am proud of my origins. Vojvodina is different from Hungary, the mentality is completely different, it's more relaxed and I feel much more at home here."

(Quote 4, Dániel)

Dorina also engages in translanguaging when she intentionally draws others' attention to her cultural and national identity, as well as her close ties to Serbia and the Balkans, which she especially highlights when commenting under videos on TikTok, where the audience is mixed. However, the participants are also aware and careful when it comes to translanguaging, as they recognize that it is still considered by many to be informal speech or text and tend to avoid it in more serious situations, especially when conversing with older individuals in positions of power (see Figure 3 above).

The anecdotes and examples (Figures 1 through 4) Dániel, Dorottya, and Dorina mention above all tell us is that their language choices and choice to engage in translanguaging are deeply rooted in and depend on their 'memberships' in specific communities or often online groups (see also Androutsopoulos, 2015; Lee, 2014). The type of consumed content and content created, and even their own beliefs about their spoken languages, as well as their confidence in those, are all additional building blocks of one's online identity (see also Lee 2014, 2016), which are carefully and intentionally curated, whether online or in face-to-face situations. Similarly to previous studies researching translanguaging and exploring the ways in which its promotion could yield positive effects especially in minority settings (Cenoz and Gorter 2017, 2019), the participants in the present study have expressed that translanguaging is something very unique to them, as they are able to understand one another on a deeper level and discuss topics that necessitate the use of all of their spoken languages for mutual understanding and effective communication. The necessity for translanguaging in such instances is something that strengthens their membership in the Vojvodina Hungarian community, where the fact that they share the same languages, a mixture of

cultures, and lived experiences with fellow Vojvodina Hungarians is something to be valued and cherished. This provides great grounds for a supportive community, where the use of their Hungarian minority language (whether through translanguaging or not) becomes a necessity, especially for discussing various topics that might not be possible without the use of Vojvodina Hungarian. Considering that Hungarian is their first language and the language they have the richest repertoire of, it is very well possible that group communication would not take place effectively if this crucial condition was not met (cf. Cenoz and Gorter, 2017, 2019).

5. Conclusion

This small-scale study has examined the digital habits and language use practices of Vojvodina Hungarians, with a specific focus on translanguaging. To explore the topic, semi-structured interviews were conducted with some of the same participants from Kostic (2024), and the findings showed that the Vojvodina Hungarian interviewees had very individual yet overlapping "life stories in relation to technology" (Lee, 2014, p. 94) as well as linguistic choices and linguistic habits in general. Although the present study is not generalizable due to its limited number of participants, the qualitative approach allowed for more detailed and individualized data collection and analysis. The findings revealed a high level of digital integration in the daily lives of the Vojvodina Hungarian participants, with a substantial amount of time spent online and a diverse range of online activities, including social media use, information seeking, communication, and various online errands. The findings also highlight the crucial role of digital platforms in maintaining social connections within the Vojvodina Hungarian community, which also fosters and provides multimodal platforms for intentional translanguaging. The internet also serves as a primary source of information for news, culture, and community-relevant information. Furthermore, the findings emphasize the multilingual nature of their communication, with frequent intentional as well as unintentional translanguaging between their spoken languages of Hungarian, Serbian, (and sometimes Croatian) and English. Their spoken and written interactions are characteristic of multilingual speech, where translanguaging plays a significant role in shaping their online identities and often occurs to facilitate more efficient communication and convey their message across as easily and effectively as possible. Their language choices are influenced by a combination of contextual, social, and personal factors, consistent with the findings of previous studies on similar topics (Androutsopoulos, 2015; Lee, 2014).

References

- Aleksić, G., and O. García. (2022). Language beyond flags: teachers' misunderstanding of translanguaging in preschools. *International Journal of Bilingual Education and Bilingualism*, 25(10): 3835-3848. Doi: 10.1080/13670050.2022.2085029.
- Androutsopoulos, J. (2007). Language choice and code-switching in German-based diasporic web forums. In: Danet, B., and S. Herring (eds). *The multilingual internet. Language, culture and communication online*. Oxford: Oxford University Press, 340-361.
- Androutsopoulos, J. (2015). Networked multilingualism: Some language practices on Facebook and their implications. *International Journal of Bilingualism*, 19(2): 185-205.
- Balla, M., S. Buljanović, and M. Ilić. (2012). Domains of Hungarian language use in Belgrade. *Jezikoslovlje*, 13(2): 565-583.
- Belmar, Guillem, and Maggie Glass. (2019). Virtual communities as breathing spaces for minority languages: Re-framing minority language use in social media. *Adeptus*, 14:1-24.
- Cenoz, J., D. Gorter. (2017). Minority languages and sustainable translanguaging: Threat or opportunity? *Journal of Multilingual and Multicultural Development*, 38(10): 901-912.
- Cenoz, J., D. Gorter. (2019). Multilingualism, translanguaging, and minority languages in SLA. *The Modern Language Journal*, 103(1): 130-135.
- Ching, C. C., and L. Vigdor. (2005). Technobiographies: Perspectives from education and the arts. Presentation, First International Congress of Qualitative Inquiry, University of Illinois Urbana-Champaign, IL, USA, May 5-7, <http://www.iiqi.org/C4QI/httpdocs/qi2005/papers/ching.pdf>.
- Đerčan, B., D. Bjelajac, and M. Bubalo-Živković. (2020). Sustainability of ethnic groups in Vojvodina, Serbia: Role of social media. *Handbook of the Changing World Language Map*. Switzerland: Springer International Publishing, 3633-3652.
- Ferdinand, S., and F. Komlosi. (2017). The use of Hungarian and Serbian in the city of Szabadka/Subotica: An Empirical Study. *Hungarian Cultural Studies*, 10: 1-13. Doi: 10.5195/ahca.2017.278.
- Galács, A., and B. Ságvári. (2013). A nyelvi kultúrák és közösségek a digitális bennszülött fiatalok körében. In: Szabó, A., B. Bauer, and P. Pillók (eds). *MOZAIK2011 – Magyar fiatalok a Kárpátmedencében*. Budapest: Belvedere Meridionale Kiadó, 359-374.
- García, O., and R. Otheguy. (2019). Plurilingualism and translanguaging: Commonalities and divergences. *International Journal of Bilingual Education and Bilingualism*, 23(1): 17-35.
- Goffman, E. (1990[1959]). *The Presentation of Self in Everyday Life*. London: Penguin.
- Hargittai, E. (2010). Digital na(t)ives? Variation in Internet skills and uses among members of the "Net Generation". *Sociological Enquiry*, 80: 92-113.
- Helsper, E. J., and R. Eynon. (2010). Digital natives: where is the evidence?. *British Educational Research Journal*, 36(3):503-520.
- Jánk, I., and Sz. Rási. (2023). Hungarian as a minority and majority language in different language policy contexts. *Current Issues in Language Planning*, 1-18.
- Jokinen, K., K. Hiovain, N. Laxström, I. Rauhala, and G. Wilcock. (2017). DigiSami and digital natives: Interaction technology for the North Sami language. *Dialogues with Social Robots: Enablements, Analyses, and Evaluation*, 3-19.
- Jongbloed-Faber, L., H. Van de Velde, C. van der Meer and E. Klinkenberg. (2016). Language Use of Frisian Bilingual Teenagers on Social Media. *Treballs de Sociolingüística Catalana*, 26: 27-54.
- Kelly-Holmes, H. (2004). An analysis of the language repertoires of students in higher education and their language choices on the Internet (Ukraine, Poland, Macedonia, Italy, France, Tanzania, Oman and Indonesia). *International Journal of Multicultural Societies*, 6(1): 29-52.
- Kornai, A. (2013). Digital Language Death. *PLoS ONE*, 8(10): e77056. [online] Doi: <https://doi.org/10.1371/journal.pone.0077056>.
- Kostic, A. (2024). Minorities in the digital space: Vojvodina Hungarians' digital presence, language choices, and language use online. *Philologia*, 22(1): 15-51. doi: <https://doi.org/10.18485/philologia.2024.22.22.2>.

- Kovacs Rac, E., and S. Halupka-Rešetar. (2018). Sense of local identity, attitudes toward dialects and language teaching: The Hungarian minority in Serbia. *Poznan Studies in Contemporary Linguistics*, 54(1): 115-146.
- Kovács Rác, E. (2012). Kisebbségi nyelvjárási attitűd és identitástudat. *Vajdasági magyar tudóstalálkozó. Tudomány, módszer, argumentáció*, 35-41.
- Lee, C. (2014). Language choice and self-presentation in social media: the case of university students in Hong Kong. In: Seargeant, P., and C. Tagg (eds). *The language of social media: Identity and community on the Internet*. 91-111.
- Lee, C. (2016). *Multilingualism Online* (1st edition). London: Routledge.
- Li, W. (2011). Moment analysis and translanguaging space: Discursive construction of identities by multilingual Chinese youth in Britain. *Journal of Pragmatics*, 43: 1222-1235.
- Lynn, T., K. Scannell and E. Maguire. (2015). Minority language Twitter: Part-of-speech tagging and analysis of Irish tweets. *Proceedings of the ACL 2015 Workshop in Noisy User-generated Text*, 1-8.
- Mandić, M., and K. Rác. (2023). Learning the language of social environment: The case of Hungarian in Vojvodina (Serbia). *Current Issues In Language Planning*, 24(4): 460-480.
- Molyneaux, H., S. O'Donnell, C. Kakekaspan, B. Walmark, P. Budka and K. Gibson. (2014). Social Media in Remote First Nations Communities. *Canadian Journal of Communication*, 39(2): 275-288.
- O'Carroll, A. D. (2013). An Analysis of How Rangatahi Māori Use Social Networking Sites. *MAI Journal*, 2(1): 46-59.
- Otheguy, R., O. García, and W. Reid. (2015). Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review*, 6(3): 281-307.
- Pásztor-Kicsi, M. (2012). Van-e vajdasági magyar nyelvhasználat a Facebookon? *Hungarológiai Közlemények*, 43(1): 42-52.
- Pásztor-Kicsi, M. (2016). Regionality–Language–Internet. *Acta Universitatis Sapientiae, Philologica*, 8(2): 35-47.
- Pauwels, A. (2016). *Language maintenance and shift*. Cambridge: Cambridge University Press.
- Pilipenko, G. (2016). Hungarian-Slavic Bilingualism in Transcarpathia, Vojvodina and Prekmurje. *Ежегодник финно-угорских исследований*, 10(4): 20-30.
- Prensky, M. (2001). Digital natives, digital immigrants. *On the horizon*, 9(5): 1-6.
- Statistical Office of the Republic of Serbia. (2012). Ethnicity. Data by Municipalities and Cities. Belgrade: Statistical Office of the Republic of Serbia.
- Szabó, A., B. Bauer and P. Pillók (eds). (2013). *MOZAIK2011 – Magyar fiatalok a Kárpátmedencében*. Budapest: Belvedere Meridionale Kiadó.
- Tagg, C. (ed). (2015). *Exploring Digital Communication: Language in action*. London: Routledge.
- Tagg, C., and P. Seargeant. (2012). Writing systems at play in Thai-English online interactions. *Writing Systems Research*, 4(2): 195-213.
- Tapscott, D. (1998). *Growing up digital*. San Francisco: McGraw-Hill Companies.
- Trombitás, T., and É. Szügyi. (2019). Education Language Choice of Hungarian ethnic diaspora communities in Vojvodina (Serbia). *Deturope*, 11(2): 54-74.
- Underwood, J. (2007). Rethinking the digital divide: Impacts on student tutor relationships. *European Journal of Education*, 42(2): 213-222.
- Vaisman, C. (2011). Performing girlhood through typographic play in Hebrew blogs. In: Thurlow, C., and K. Mroczek (eds). *Digital discourse: Language in new media*. Oxford, UK: Oxford University Press, 177-196.

Appendix: Interview Questions

A. Background information (age, gender, place of birth, current place (settlement) of residence, occupation, spoken languages, mother tongue, nationality)

B. Linguistic habits, attitudes, and spoken languages

- Do you use Serbian/English in your daily life, and have you encountered situations where knowledge of Serbian/English is important? Do you feel pressure to become proficient in Serbian/English? If so, for what reasons?
- How would you describe your relationship with Hungarian, Serbian, and English?
- How do you feel about mixing languages in conversations (online and face-to-face)? Is it common among people you know? Can you share some examples of situations where you or others mix languages in conversation?
- Are there specific situations where you prefer one language over another? Why?

C. Digital habits and language use online

- What languages do you primarily use when communicating online? Do you find it easier to express yourself in one language over another on social media? For what reasons?

- Are there any challenges or benefits you associate with using multiple languages digitally?

- Can you describe what your internet habits and device usage used to look like 10-15-20 years ago when the internet was a novelty in our region? What about nowadays, how and what for do you use the internet and your devices?

- How frequently do you consume and create digital media, such as news articles, videos, photos, or even podcasts? Are there specific social media sources that you trust or rely on more than others? If so, for what reasons?

D. Online communities in Vojvodina

- How do you experience the digital space as a Vojvodina Hungarian? Have you encountered discrimination or challenges related to your national and/or linguistic identity online?
- Are there online communities or resources that have helped you connect with your cultural or linguistic heritage? Is it important to you to preserve and promote the use of the Hungarian language in this region?

WAKWOYA MIHIRETU¹ – MARGARET DEUCHAR²

¹University of Pannonia
tolosa.mihiretu@phd.uni-pannon.hu
ORCID id:<https://orcid.org/0000-0003-4001-9139>

² University of Pannonia
md118@cam.ac.uk

Wakwoya Mihiretu – Margaret Deuchar: Analyzing Code-Switching Between Afaan Oromoo and English in Ethiopia: A Grammatical Perspective
Alkalmazott Nyelvtudomány, XXV. évfolyam, 2025/1. szám, 104–121.
doi:<http://dx.doi.org/10.18460/ANY.2025.1.006>

Analyzing Code-Switching Between Afaan Oromoo and English in Ethiopia: A Grammatical Perspective

This pilot study investigates code-switching between Afaan Oromoo and English in informal community interactions in Dambi Dollo, Ethiopia, using the Matrix Language Frame (MLF) model. It is based on two audio recordings totaling 90 minutes of spontaneous conversations, which have been transcribed and analyzed. The MLF model's principles were applied to identify the Matrix Language (ML) and Embedded Language (EL). Clauses are classified into monolingual and bilingual, with Afaan Oromoo predominantly governing syntactic structure, while English mainly serves as a lexical provider. The analysis highlights a strong preference for Afaan Oromoo in informal conversations, with Amharic and English playing minor roles. These findings contribute to understanding multilingualism in Ethiopia, demonstrating how linguistic structures and social factors influence bilingual communication.

Keywords: Afaan Oromoo, Bilingualism, Code-switching, English, Matrix Language Frame (MLF) Model

1. Introduction

Ethiopia, with its diverse communities, each distinguished by distinct languages, cultures, and ethnicities, offers a valuable opportunity for studying bilingualism and multilingualism in everyday life. According to Meyer et al. (2023), Ethiopia, an East African nation with a population exceeding 100 million, is home to over 80 languages, although this number can vary depending on the methodology used. For instance, Ado et al. (2021) identified over 85 languages spoken by approximately 110 million people. The coexistence of so many languages in Ethiopia creates a unique environment for studying various linguistic phenomena, with code-switching (CS) being a particularly prominent feature.

CS is a linguistic phenomenon observed in bilingual speech or writing, where individuals use two or more languages within a single conversation or piece of discourse (Deuchar et al., 2018; Deuchar & Stammers, 2012). In this context, CS refers to the practice of mixing multiple languages within the same communicative

exchange. This phenomenon is widespread in informal settings and multilingual communities, such as Ethiopia, where individuals fluidly mix linguistic elements to convey meaning and express cultural nuances. See instances of code-switching in the following examples. Key to glosses: 1/2/3PL, First/second/third Person Plural; 1/2/3SL, First/second/third Person Singular; POSS, possessive pronoun; DET, determiner; ACCO, Accusative; Q, Question mark; ART, Article; NEG, negative/negative particle; IMPV, imperfect verb; IMP, Imperative verb; PRV, Perfective Verb; CNV, Converb; FOC, Focus marker; COP, copula; NOM, Nominative marker; LOC, Locative marker; M, male; F, female; Pass, passive marker. Additionally, Afaan Oromoo words appear in standard font, English words in **bold** with @eng, and Amharic words in *italics* with @amh.

(1) **Direct** @eng hin did-e maalinnii wanti-chi.
Direct NEG refuse-3SL.PRV. what thing-ART
 “It refused to direct what the thing is.”
 (Maccaa-OC09-MAB-94)

(2) *Mastaaweqiiyaa*@amh godh-ee achi kaa’-e
Advertisement do-3SL.M.CNV.PRV there put-3SL.M.PRV
 factory
 factory

blue magic@eng-tu.
blue magic -FOC
 “The Blue Magic factory placed it there as an advertisement.”
 (Maccaa-OC09-SAF-280)

(3) Kun **file**@eng -tti hidh-am-a.
 This **file**- to tie-PASS-IMPV
 “This is tied to a file.”
 (Maccaa-OC09-SAF-48)

(4) **Photo**@eng hin qab-uum
photo NEG have-3SL.CNV.IMPV
 “It does not have a photo.”
 (Maccaa-OC09-MAB-45)

CS has been extensively researched from various perspectives, employing diverse approaches and theoretical frameworks. Researchers have examined CS through linguistic, sociocultural, cognitive, and psycholinguistic lenses, as demonstrated in the works of scholars (such as Auer, 2013; Bullock & Toribio, 2009; Deuchar & Stammers, 2012; Gardner-Chloros, 2009); Khan & Khalid, 2018; Wei, 2009; and are among others. However, there is a notable lack of studies on CS within the Ethiopian context. Some Ethiopian research has addressed this topic (see Ali 2015; Bejiga 2021; Leyew 1998; and Sime 2019, for more details), but the Ethiopian context remains underexplored compared to studies in Western countries. Consequently, the study to be reported here is of potential significance. There is a gap in the literature regarding code-switching between Afaan Oromoo and other languages from a linguistic perspective in Ethiopia.

Sime (2019) explores Amharic-English code-switching in Ethiopian EFL classrooms, comparing its frequency and types in primary (grade 7) and secondary (grade 9) levels. Findings show higher CS use in primary classrooms (31.9%) compared to secondary classrooms (17%). Four CS types are identified: inter-sentential, intra-sentential, extra-sentential (tag), and intra-word switching, with intra-sentential CS being dominant at the primary level and inter-sentential at the secondary level. The study concludes that while CS supports learning, its use should align with students' proficiency to ensure adequate English exposure.

Keleta (2020) investigated code-switching between Tigrinya and English within FM radio broadcasts in Mekelle, Ethiopia, using the Matrix Language Frame (MLF) model. This research examines intrasentential code-switching in the Tigrinya language, with a focus on the prevalence of English nouns, adjectives, and verbs in radio broadcasts. Keleta's study highlights the role of Tigrinya as the matrix language, with English elements fitting into the morphosyntactic structure of Tigrinya (Keleta, 2020).

In a similar vein, Emam & Mekonnen (2022) examined code-switching between Amharic and English within the Ethiopian media. They found that English elements such as nouns, adjectives, and adverbs were frequently embedded within Amharic sentences, with Amharic serving as the matrix language. Their study underscores the dynamic relationship between Amharic and English in Ethiopia's multilingual media landscape and highlights the influence of English in modern civic and technological discourse (Emam & Mekonnen, 2022).

Leyew (1998) investigates Amharic-English code-switching in Ethiopia, analyzing linguistic and sociolinguistic factors across various contexts, including schools, universities, and media. The study identifies Amharic as the matrix language and English as the embedded language, with nouns and adjectives being switched more frequently than verbs due to morphological limitations. Influenced by

education, prestige, and linguistic economy, code-switching is systematic rather than random, though it is often viewed negatively by monolinguals. The study highlights structured patterns in Amharic-English bilingual interactions.

The primary aim of this pilot study, to be reported, was to explore code-switching (CS) between Afaan Oromoo and English in informal community interactions, where bilingualism and multilingualism are prevalent. The Matrix Language Frame (MLF) theory will provide the analytical framework (Myers-Scotton, 1993, 2002, 2004, 2006). The MLF distinguishes between the Matrix Language (ML) and the Embedded Language (EL) in code-switched clauses. The analysis will focus on clause-level code-switching, specifically intraclausal code-switching, which involves switching within a clause, as opposed to interclausal code-switching, switching that occurs at clause boundaries (Deuchar, 2012). This study will examine the distribution of CS between Afaan Oromoo and English, with particular attention to identifying the Matrix Language (ML).

2. Literature Review

2.1 Previous Research on Code-Switching in Ethiopian Languages

Despite Ethiopia being a multilingual country, there is a general scarcity of research exploring code-switching between pairs of languages, with little research on code-switching between Afaan Oromoo and English. I have found only a limited number of studies on Ethiopian languages such as English /Amharic, English/Tigrigna and rarely Afaan Oromoo/English CS, and these were in educational settings and media. Among the Ethiopian studies reviewed in Section 1, there are relevant insights into this area (see Balay, 2020; Bejiga, 2021; Leyew, 1998; and Sime, 2019, for details). In addition to these local studies, Temesgen and Hailu (2022) and Ali (2015) also investigated CS between Amharic/English and Oromiffa-Harari languages, respectively.

Temesgen and Hailu (2022) investigate code-switching practices among EFL teachers in Ethiopia, addressing a research gap in multilingual education contexts. Their study reveals that teachers switch between Amharic and English to clarify complex concepts, explain vocabulary, manage classroom dynamics, and foster rapport. Code-switching is driven by students' limited English proficiency, teachers' instructional beliefs, and the specific language skills being taught. The authors argue for the strategic use of code-switching as a pedagogical tool, rather than enforcing an English-only approach (Temesgen & Hailu, 2022).

Ali (2015) investigates Oromiffa-Harari code-switching in Dire Dawa, highlighting both its structural patterns and sociolinguistic motivations. The study, based on interviews, focus groups, and recorded conversations, finds that Oromiffa functions as the matrix language, with Harari as the embedded language. Code-

switching occurs systematically across word categories such as nouns, verbs, and adjectives, with intra-sentential, inter-sentential, and tag-switching being the most common types. Motivation includes ethnic identity expression, communicative efficiency, and bilingual competence. Notably, older speakers tend to use inter-sentential switching, while younger ones prefer intra-sentential forms. While offering valuable insights into bilingualism involving Afaan Oromoo, the study also underscores a research gap—there is limited investigation into clause-level Afaan Oromoo-English code-switching in informal, everyday discourse.

2.2. Theoretical Framework: The Matrix Language Frame Model

The Matrix Language Frame (MLF) model, developed by Myers-Scotton, is a key theoretical framework in the study of code-switching. It distinguishes between the Matrix Language (ML), which determines the syntactic structure of a clause, and the Embedded Language (EL), which contributes lexical items. A key idea of the model is that while the ML sets up the grammar, the EL adds words in a way that follows the ML's rules without altering them (Myers-Scotton, 1993, 2002).

The MLF model has been widely used to analyze code-switching in various bilingual communities (Deuchar et al., 2018). The key principles of the MLF model include the Matrix Language Principle, Morpheme Order Principle (MOP), and System Morpheme Principle (SMP). These principles guide the identification of matrix language by analyzing word order and grammatical markers (Deuchar, 2006).

The 4-M refined the MLF model by categorizing morphemes into content and system morphemes. System morphemes are further divided into early system morphemes, bridge late system morphemes, and outsider late system morphemes. I will employ the updated version of the MLF Model in my study to analyze the structural and morphosyntactic patterns of code-switching between Afaan Oromoo and English (see Jake, J. L., & Myers-Scotton, 2020; Myers-Scotton & Jake, 2017; Myers-Scotton & Jake, 2015; Myers-Scotton, 2002, for more details).

3. Methodology

3.1 Data Collection

The data for this study were drawn from a bilingual corpus consisting of conversations in Afaan Oromoo-English. Many of the speakers were also competent in Amharic, one of the official languages. These conversations were recorded with native speakers of Afaan Oromoo, transcribed, and systematically coded to identify instances of code-switching. To maintain consistency in the transcription of words from two or more languages, the following conventions were employed: words in Afaan Oromoo were rendered in standard font, English words were presented in bold and tagged as @eng, while Amharic words were italicised and tagged as @amh. Each

example was annotated with detailed morphological information, including grammatical features such as tense, aspect, and number.

In this study, transcription files are named to ensure consistency, facilitate efficient data management, and traceability throughout the analysis. Each filename encapsulates key metadata, including dialect type, recording sequence, anonymized speaker identity, and the precise location of the extracted utterance within the dataset. In the file name Maccaa-OC09-MAB-45, “Maccaa” shows the specific Afaan Oromoo dialect; ‘OC’ stands for the dataset (with ‘O’ for Oromo and ‘C’ as a special code); “09” is the recording's sequence number; “MAB” is a fake name to keep the speaker's identity private; and “45” points to the row in the spreadsheet where this example is found.

This study draws on two audio recordings comprising approximately 90 minutes of spontaneous conversation among four speakers. We transcribed and organized the data into a structured spreadsheet. We identify each speaker using pseudonyms and delineate their corresponding utterances throughout the dataset. Utterances are classified as monolingual or bilingual clauses. The spreadsheet provides interlinear glossing for each clause to illustrate the underlying morphological and syntactic structures, with English translations for clarity.

The study employed a social network approach to recruit participants (Milroy, 1987). This ensures a diverse sample representing various social and occupational backgrounds. Speakers, aged 20 to 40, included government employees and students. All participants spoke the Macca Oromo dialect of Afaan Oromoo and had lived in the Kellem Wollega Zone for over 20 years.

3.2 Linguistic Data Analysis

The analysis aimed to identify the role of each language (Afaan Oromoo and English) as either the Matrix Language (ML) or Embedded Language (EL). Using the MLF model, we looked at the structure of each sentence to see if it followed the usual SOV (Subject-Object-Verb) order of Afaan Oromoo or the SVO (Subject-Verb-Object) order of English.

The data analysis process categorized utterances into simple and complex clauses. Each identified clause functioned as a unit of analysis to determine the Matrix language, as outlined above. In this analysis, main clauses and subordinate clauses are treated as distinct units. The clauses were classified as monolingual and bilingual on an Excel spreadsheet. Each clause was glossed, translated into English, and labeled as follows: monolingual Afaan Oromoo, monolingual English, monolingual Amharic, and bilingual clauses. If a clause has morphemes from one language, it's monolingual; if it has morphemes from two or more, it's bilingual.

This study seeks to identify the matrix language within the context of code-switching between Afaan Oromoo and English. We analyzed word order, constituent structure, and finite verbs to determine the Matrix Language (ML) in Oromo/English clauses based on specific criteria, including morpheme order and outsider late morpheme principles. The assumption is that the morphology of the finite verb, i.e., outside late system morphemes, will originate from the ML if the clause shows classic code-switching patterns.

3.3 Distinguishing Code-Switching from Loanwords

Establishing clear criteria for differentiating between lexical loans and single-word switches in the code-switching between Afaan Oromoo and English is crucial for accurately analyzing language contact. This also offers insights into how external languages have influenced the structure and vocabulary of Afaan Oromoo over time.

In this study, words in Afaan Oromoo that were listed in authoritative dictionaries were considered loans, while words not in the dictionary were categorized as code-switches. This approach helps to clarify the relationship between language mixing and the incorporation of foreign elements in bilingual speech.

We have included single-word switches and differentiated them from loanwords based on their predictability (Muysken, 2000, p. 71; cf. Deuchar, 2006, p. 1899). This predictability is linked to “listedness,” which refers to whether a word is stored in the speaker’s mental lexicon, like being listed in a dictionary. We relied on established dictionaries, particularly the “Elelle trilingual Afaan Oromoo Amharic and English Dictionary” (Hinsen Makuria, 2009), to assess listedness. If an English-origin term is listed in a recognized Afaan Oromoo dictionary, it is classified as a loan; if absent, it is treated as a switch. We recognized that this approach may not be watertight, as dictionaries may not fully reflect current usage, which can lead to potential misclassifications. In the context of this study, which aims to identify the matrix language in code-switching instances between Afaan Oromoo and English, loanwords are excluded from the analysis. See the following examples:

(5) Kaleessa	akkuma	ati	jetteen	cufe
Yesterday	as	you	say-2SL.CNV.PRV	shut-1SL.PRV
moobaayilii				
mobile				

“Yesterday as you said I shut down the mobile.”

(Maccaa-OC09-SAF-115)

- (6) **Form@eng** gaafa guutte moo bilbileen itti
 Form when fill-3SL.F.PRIV while call-1SL.PRIV to
 him-a Yoonammoo
 tell- 1SL.IMPV Yona-FOC
 “When you fill out the form, I will call and tell them now.”
 (Maccaa-OC09-MAB-289)

In example (5), the Afaan Oromo dictionary lists the term “moobaayilii” (mobile) as a loanword from the English language. On the other hand, in example (6), the word ‘form’ is not listed in the “Elelle Afaan Oromoo dictionary” and “Glosbe” online Oromo-English dictionary, and so is categorized as the switch to English (the equivalent meaning in Afaan Oromoo is “unka” (form)).

4. Results

4.1 Distribution of Matrix Language

4.1.1 Monolingual Clauses

The ML model framework applies not only to bilingual clauses but also to monolingual ones. Specifically, the Morpheme Order Principle and System Morpheme Principle can be used to find matrix language within these monolingual contexts, as well as in bilingual contexts. Afaan Oromoo follows a Subject-Object-Verb (SOV) order in declarative sentences, while English follows a Subject-Verb-Object (SVO) order. See the following example.

- (7) Nam-ni amantii isaa sodaat-a. (Maccaa-OC01 EYN 544)
 Man-NOM religion his fear-IMPV
 “Man fears his religion.”

In example (7), the clause follows the typical Subject-Object-Verb (SOV) structure of Afaan Oromoo, with the subject “Nam-ni” marked by the nominative case “-ni,” the object “amantii isaa” (his religion) with the possessive marker “isaa,” (his), and the verb “sodaat-a” (to fear) in its finite form. The verb ends with the ‘-a’ aspect marker, a late outsider morpheme in Afaan Oromoo, indicating the imperfective aspect. This structure confirms Afaan Oromoo as the matrix language, with the “-a” aspect marker aligning with its grammatical system. The following examples will be used to illustrate the application of the system Morpheme Principle.

- (8) Lafa meeqa bitt-a. (Maccaa-OC01 DAH 220)
 Land much buy-2S. IMPV
 “You buy much land”

In example (8), the analysis confirms Afaan Oromoo as the matrix language.

The clause follows the typical SOV structure, but the object “lafa” precedes the verb “bitta” due to the implied subject. The adjective “meeqa” (a lot/much) follows the noun ‘lafa,’ consistent with Afaan Oromoo’s noun-adjective order. The verb ‘bitt’ is marked with the -a suffix, a late outsider morpheme that indicates the 2nd person singular subject “you” and the imperfective aspect, signaling a habitual action. This structure highlights the verb morphology and syntax typical of Afaan Oromoo.

No English monolingual clauses were found in the two conversations analyzed in detail for this study. However, example (9), drawn from a separate recording (Maccaa-OC02-SAG 489) that was not closely examined, is included as an instance to illustrate that such clauses may appear in the larger dataset.

- (9) **He never gives up.**
He never give-3SL.M. IMPV up.
 “Inni gonkumaa abdii hin kutatu.”(equivalent Afan Oromoo translation)

This clause follows the English SV order and satisfies the System Morpheme Principle, confirming English as the matrix language. The finite verb “gives” agrees with the subject “He” and demonstrates a late outsider morpheme through third-person singular agreement. The pronoun “He” is a content morpheme, while the adverb ‘never’ is an early system morpheme modifying the verb. The particle ‘up’ functions as a bridge late system morpheme, linking with “gives” to form the phrasal verb “gives up.” These elements, such as the Morpheme Order Principle and the 4-M model, confirm English as the matrix language. The following example illustrates the syntactic structure of Amharic, a Semitic language with SOV word order in declarative clauses that employs the Ge’ez (Fidel) script, an alpha syllabary (abugida) derived from the ancient Ge’ez writing system. The following Amharic example demonstrates these features.

- (10) ፀጉራ ደረቀ ተበጣጠሰ፡፡ (Maccaa-OC09 SAF 323)
 Ts’egur-ē derek’-e tebet’at’es-e.
 Hair-my dry-PFV tangled-PFV
 ‘My hair dried and got tangled.’

In example (10), the clause deviates from the typical SOV structure of Amharic, following a Subject-Verb-Verb (SVV) sequence instead, with the subject $\theta\eta\zeta$ (Ts'egurē, 'my hair') marked by the possessive suffix "–ē". The two perfective verbs $\mathfrak{L}\mathfrak{L}\Phi$ (derek'e, 'dried') and $\mathfrak{t}\mathfrak{n}\mathfrak{r}\mathfrak{r}\mathfrak{h}$ (tebet'at'ese, 'became tangled') suggest a causal relationship. Although the clause lacks an explicit object and does not fully align with the Morpheme Order Principle, the 4-M Model helps identify Amharic as the Matrix Language. The content morphemes— $\theta\eta\zeta$ (Ts'egurē), $\mathfrak{L}\mathfrak{L}\Phi$ (derek'e), and $\mathfrak{t}\mathfrak{n}\mathfrak{r}\mathfrak{r}\mathfrak{h}$ (tebet'at'ese)—hold the primary meaning, while the possessive morpheme '-ē' and the perfective aspect markers '-e' are seen as early and late system morphemes, respectively. All these morphemes are derived from Amharic's morphological system, confirming Amharic as the matrix language in this clause.

4.1.2. Bilingual Clauses

- (11) Utuu **initiative**@eng erga ta'-ee Qeellam Wallaggaa -tti
If initiative since be-CNV-PRV Kellem Wallagga -to
boqolloo -tu ta'-a.
corn -FOC be-IMP
“If there were an initiative, there would be corn in Kellem Wallagga.”
(Maccaa-OC01 DAH 21)

“boqolloo” (corn) as a content morpheme, “-tu” as an early system morpheme (focus marker), and “ta’-a” as the finite verb marked for modality and agreement (late outsider morpheme). The presence of these system morphemes from Afaan Oromoo confirms it as the matrix language.

- (12) **Percent@eng** jaatamni kun eessa dhaq-aa ?
 Percent sixty-NOM this where go-CNV.IMPV -Q
 “Where does this sixty percent go?”
 (Maccaa-OC01 EYN 139)

Example (12) follows the Subject–Interrogative Word–Verb (SIV) order typical of Afaan Oromoo interrogatives. The subject phrase “Percent jaatamni kun” (this sixty percent) reflects Afaan Oromoo noun phrase structure, with the head noun (‘percent’) first, followed by the numeral “jaatamni” and demonstrative “kun”, contrasting with English word order. The interrogative “eessa” (where) follows the subject, and the verb “dhaqaa” (go) completes the clause, adhering to Afaan Oromoo syntax. Though “percent” is an English content morpheme, all system morphemes—including “jaatamni,” “kun,” “eessa,” and the imperfective aspect marker “-aa” which is a late outsider morpheme—are from Afaan Oromoo. This confirms Afaan Oromoo as the matrix language governing the clause's grammatical structure.

- (13) Nageenya jech -uu-n **business@eng** ta’-e.
 Security mean -NP business be-3SL.PRV
 “Security became a business.”
 (Maccaa-OC01 EYN 190)

In Example (13), the clause adheres to the morphosyntactic structure of Afaan Oromoo and includes an English insertion. “Nageenya” (security) is the subject, “jechuun” (means) functions as a noun/verb hybrid, and “business” is the English noun insertion. “ta’e” (became) is the finite verb with a perfective marker “-e.” Despite the insertion, the clause adheres to Afaan Oromoo grammar. In Myers-Scotton’s 4-M Model, “Nageenya” and “business” are content morphemes, while “jechuun” is an early system morpheme, and “-e” is a late outsider system morpheme marking tense and aspect.

- (14) **So@eng** gaafa dhimmi keenyatti deebin-uu
So, when issue our-LOC go back. 1PL.CONV. IMPV
 “So, when we go back to our issue”

(Macca-OC02-SAG 41)

In example (14), all morphemes are derived from Afaan Oromoo, except for the conjunction “so,” which is sourced from English. The finite verb “deebin-uu” (go back) agrees with the subject, which is shown by a first-person plural marking on the verb functioning as a late outsider morpheme.

- (15) **Daily @eng** hin -qaam-u
Daily NEG -chew.1SL.IMP
 “I do not chew ‘chat’ **daily**.”

(Maccaa-CO04-GAE 667)

Similarly, example (15) features the morpheme “daily” from English, while the rest of the clause is in Afaan Oromoo, and the verb “qaam-u” (chew) originates from Afaan Oromoo. Here, the finite verb “chew” appears at the end of the clause, reflecting the order, with the implied subject “I” and the suffix “-u” marks it as a late outsider morpheme.

- (16) **Attendance@eng** guut-aa-n jir-a.
Attendance fill.CONV.1SL exist-IMPV
 ‘I am filling out attendance.’

(Maccaa-OC016 SIB 96)

Also, example (16) features the morpheme “attendance” from English, while the rest of the clause is in Afaan Oromoo, and the verb “jir-a”(exist) originate from Afaan Oromoo. The finite verb “*jir-a*,” marked by the suffix “-a,” functions as an outsider system morpheme and appears after the English content morpheme “*attendance*,” which serves as the object in the clause.

- (17) Amma ijoollee **batch@eng** keenyaa yoo gaafat-t-ee
 Now children **batch** our if ask.2P-CONV.PRIV
 “Now, if you ask the students of our batch.”

(Maccaa-OC016 BOO-307)

In example (17), the clause adheres to Afaan Oromoo syntax while incorporating the English word “batch.” The phrase “ijoollee batch keenyaa” follows a noun–modifier–possessive structure, contrasting with English order. The verb “gaaffattee” (you asked) reflects typical SOV order, with the suffix “-ee” functioning as a late outsider morpheme. Despite the insertion, the clause aligns with Afaan Oromoo grammar, confirming it as the matrix language.

According to Myers-Scotton's Matrix Language Frame Model, the Morpheme Order Principle (MOP) and system morpheme principles assert that the morpheme order in examples (7), (8), (11), (12), (13), (14), (15), (16), and (17) reveals that Afaan Oromoo is the matrix language in all of these examples. English often limits its influence on content words such as nouns and verbs, while the grammatical framework conforms to Afaan Oromoo principles. Example (18) further illustrates this pattern, as the syntactic structure of Afaan Oromoo maintains its dominance even with the insertion of an English lexical item.

- (18) Xaafii nam-ni **export@eng** hin -godh-u.
Teff.ACCO man.NOM **export** NEG do.IMPV
“Man does not **export** teff”

(Maccaa-OC12 EYS-326)

In example (18), the clause is primarily in Afaan Oromoo, with the English insertion “export.” The object “Teff” precedes the subject “Man,” forming an OSV order, atypical for Afaan Oromoo and possibly pragmatically motivated. The final verb, “hin godhu” (do not), aligns with typical SOV syntax, reinforcing Afaan Oromoo as the matrix language.

- (19) *Biri@amh* dhibba lama **transfer@eng** naa godh-i.
Birr hundred two **transfer** **for** me do-IMP.IMPV
“Do **transfer** two hundred birrs to me.”

(Maccaa OC010 YOT 350)

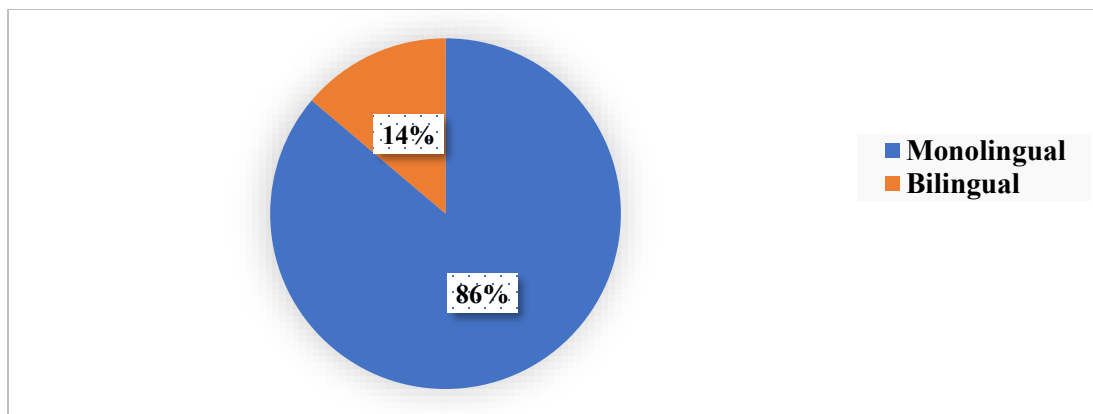
In example (19), three languages are represented: “Birr” is from Amharic, “transfer” is from English, and the remaining terms are from Afaan Oromoo. Despite the insertions, the clause maintains Afaan Oromoo's grammatical structure, highlighting its role as the matrix language in this trilingual code-switching instance. The phrase “Birr dhibba lama” (‘two hundred Birr’) serves as the direct object of the imperative verb “godhi” (‘do/make’), while “naa” indicates the indirect object (to me), with the subject “you” implied, as is typical in Afaan Oromoo imperatives. Even though both Afaan Oromoo and Amharic use an SOV structure, the System

Morpheme Principle indicates that Afaan Oromoo is the matrix language because all the system morphemes come from it, with just one Amharic content morpheme included. This illustrates how the clause maintains the grammatical framework of Afaan Oromoo while accommodating lexical insertions from another language.

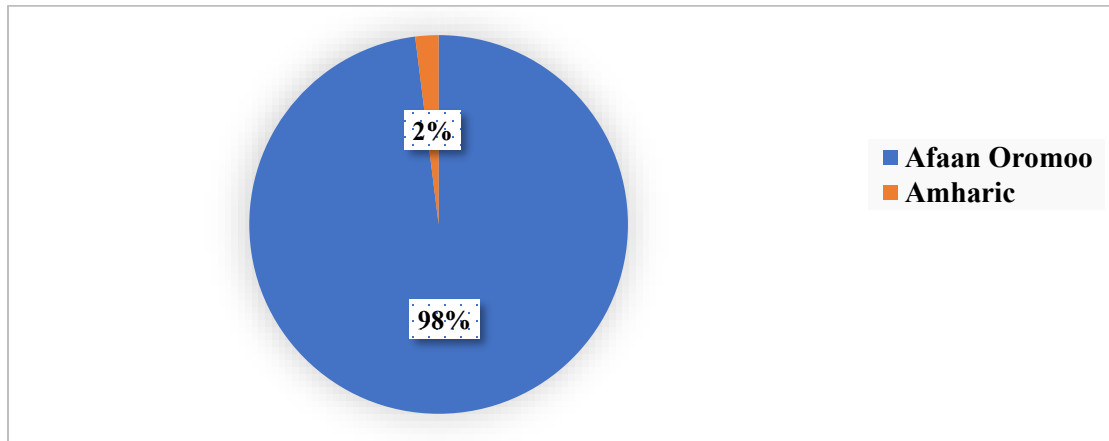
4.2 Quantitative Analysis

The transcription analysis yielded a total of 644 clearly interpretable clauses, comprising 555 monolingual and 89 bilingual clauses. Clauses deemed unintelligible—due to factors like poor audio quality, overlapping speech, or incomplete utterances—were excluded to ensure data reliability. The analyzed clauses were drawn from two distinct conversational contexts: Maccaa-OC01, where speakers aged 23 and 33 discuss market inflation and its impacts, and Maccaa-OC09, involving individuals aged 24 and 40 conversing in a human resources office during routine work.

Figure 1. The distribution of overall Monolingual and Bilingual Clauses



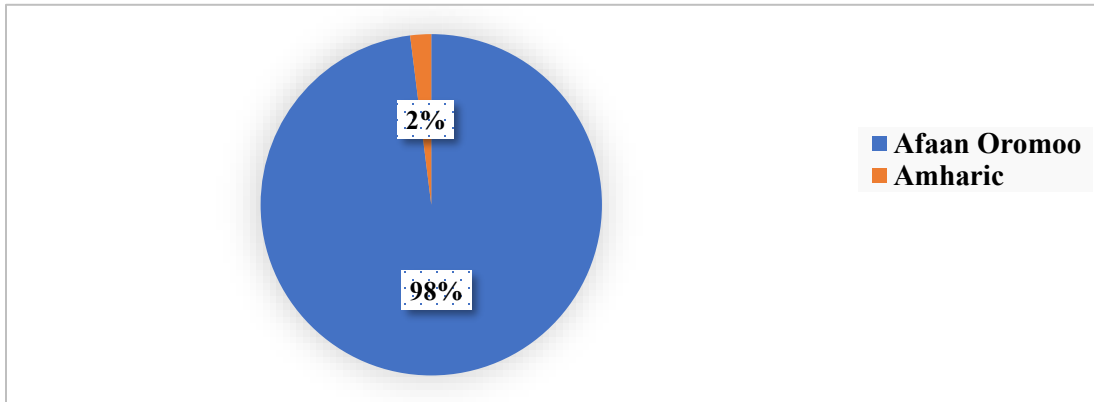
We further classified the monolingual clauses based on the conversation's language (see Figure 2).

Figure 2. Matrix Language Distribution of Monolingual Clauses

In Figure 2, of the 555 monolingual clauses, the majority comprised 544 clauses (98%) in Afaan Oromoo. This analysis highlights the predominant role of Afaan Oromoo in conversation. Eleven (2%) of the remaining monolingual clauses contained Amharic, indicating its minimal use in the discourse. There were no monolingual English clauses.

Out of the total dataset, 89 clauses—approximately 14%—were bilingual, revealing distinct patterns of language mixing. Among these, 29 clauses featured English insertions, 55 included Amharic insertions, and 5 contained both Amharic and English insertions. This distribution underscores the dominance of Afaan Oromoo across the data while also highlighting Amharic’s stronger presence than English within bilingual clauses. The higher frequency of Amharic insertions suggests its deeper integration in bilingual interactions, likely due to historical ties. This includes the geographic proximity between Amharic and Afaan Oromoo. Additionally, the occurrence of clauses with both Amharic and English insertions points to complex language contact dynamics within the speech community.

Although English was used less frequently compared to Amharic, this finding aligns with earlier research that indicates English is typically spoken in formal, global, or educational settings, whereas Amharic is more commonly used in everyday conversations between bilingual speakers.

Figure 3. Matrix Language Distribution of Bilingual Clauses

Although Afaan Oromoo predominantly functions as the matrix language in bilingual clauses, the analysis reveals two instances where Amharic assumes that role and Afaan Oromoo as the embedded language. Notably, no bilingual clauses were found with English as the matrix language. However, due to the limited dataset—drawn from only two conversations—this should not be viewed as conclusive. Although English is often used in many areas, especially in formal or international settings, its limited use as the primary language here suggests that Afaan Oromoo and Amharic play a more significant role in these specific conversations. Nonetheless, the presence of monolingual English clauses in other recordings indicates that English still holds significant discourse value within the broader communicative landscape.

These findings underscore the linguistic patterns of the speakers, with Afaan Oromoo serving as the primary language of communication, while the use of Amharic and English remained peripheral.

5. Discussion

This research provides valuable insights into code-switching between Afaan Oromoo and English in informal community interactions in Dambi Dollo, Ethiopia. The analysis shows that Afaan Oromoo overwhelmingly functions as the matrix language in bilingual discourse, with English frequently appearing as a source of lexical provider.

The analysis revealed that speakers maintained the grammatical integrity of Afaan Oromoo, even when inserting English words. This finding is based on the Morpheme Order Principle and the System Morpheme Principle from the MLF model, which state that the main language controls the sentence structure and basic parts of speech in bilingual conversations. Although English and Amharic words were included in

Afaan Oromoo sentences, they did not alter the main grammar of the conversation, indicating that Afaan Oromoo remained the primary language used.

The numbers show that most of the speech is in one language (86%), while speech that mixes languages makes up a smaller part (14%), mainly by adding Amharic and English words into Afaan Oromoo conversations, indicating that even though people switch languages, they mostly stick to using Afaan Oromoo.

6. Conclusion

This study provides significant contributions to understanding code-switching between Afaan Oromoo and English in informal community settings in Dambi Dollo, Ethiopia. The results show that Afaan Oromoo is the primary language used in conversations, and the way sentences are structured follows Afaan Oromoo rules, even when English words are used. The study reinforces the Matrix Language Frame (MLF) model, which explains how bilingual speakers maintain the grammatical integrity of one language while introducing elements from a second language.

Future research could extend these findings by examining larger datasets and considering how social variables such as education, occupation, and social networks influence code-switching patterns. Additionally, it would be valuable to investigate how cognitive processes enable bilingual speakers to navigate multiple languages without disrupting the syntactic structure of their matrix language, thereby further advancing our understanding of the cognitive and social mechanisms behind bilingual language use.

References

- Ado, D., Gelagay, A. W., & Johannessen, J. B. (2021). The languages of Ethiopia. *Grammatical and Sociolinguistic Aspects of Ethiopian Languages*, 48, 1.
- Auer, P. (2013). *Code-switching in conversation: Language, interaction and identity*. Routledge.
- Balay, B.-E. (2020). Code Switching from Amharic to English in Bilinguals. *Accessed: May, 29*.
- Bejiga, M. (2021). EFL Teachers' Perception and Practices of Code-Switching to Amharic. *Language in India*, 21(8).
- Bullock, B. E., & Toribio, A. J. E. (2009). *The Cambridge Handbook of Linguistic Code-Switching*. Cambridge University Press.
- Deuchar, M. (2012). Code switching. *The Encyclopedia of Applied Linguistics*.
- Deuchar, M. (2006). Welsh-English code-switching and the Matrix Language Frame model. *Lingua*, 116(11), 1986–2011.
- Deuchar, M., Donnelly, K., & Webb-Davies, P. (2018). *Building and using the Siarad Corpus*.
- Deuchar, M., & Stammers, J. R. (2012). What is the “nonce borrowing hypothesis” anyway? *Bilingualism: Language and Cognition*, 15(3), 649–650.
- Emam, E. H., & Mekonnen, A. M. (2022). Patterns of Code-switching in the Amharic Media. *Macrolinguistics, Vol.10 No.2 (Serial No.17)*, pp.125-150. <https://doi.org/10.26478/ja2022.10.17.6>
- Gardner-Chloros, P. (2009). *Sociolinguistic factors in code-switching* in Bullock, B. E., & Toribio, A. J. E. (2009). *The Cambridge handbook of linguistic code-switching*. Cambridge university press.
- Hinsin Makuria (2009). *Elellee. English-Oromo- Amharic Dictionary*. Addis Ababa: Graphic Printers.

- Jake, J. L., & Myers-Scotton, C. (2020). The 4-M model: Different routes in production for different morphemes. *The Routledge Handbook of Language Contact*, 63–87.
- Keleta, B. A. (2020). English-Tigrinya Intrasentential Code-switching on Tigrinya FM Radio Programs (TFRPs). *Journal of Literature, Languages and Linguistics*. <https://api.semanticscholar.org/CorpusID:219480353>
- Khan, A. A., & Khalid, A. (2018). Pashto-English codeswitching: Testing the morphosyntactic constraints of the MLF model. *Lingua*, 201, 78–91.
- Leyew, Z. (1998). Code-Switching: Amharic-English. *Journal of African Cultural Studies*, 11(2), 197–216.
- Meyer, R., Wakjira, B., & Leyew, Z. (2023). *The Oxford Handbook of Ethiopian Languages*. Oxford University Press.
- Milroy, L. (1987). Observing and analyzing natural language: A critical account of sociolinguistic methods. *Oxford: Blackwell*.
- Muysken, P. (2000). *Bilingual Speech, Typology of Code-Mixing*. Cambridge University Press. Cambridge.
- Myers-Scotton, C. (2006). *Multiple Voices: An Introduction to Multilingualism*. Blackwell. Malden, Mass.
- Myers-Scotton, C. (2004). *How code-switching as an available option empowers bilinguals*. LAUD.
- Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford University Press, USA.
- Myers-Scotton, C. (1993). Common and uncommon ground: Social and structural factors in codeswitching. *Language in Society*, 22(4), 475–503.
- Myers-Scotton, C. M., & Jake, J. L. (2017). Revisiting the 4-M model: Code-switching and morpheme election at the abstract level. *International Journal of Bilingualism*, 21(3), 340–366.
- Myers-Scotton, C., & Jake, J. L. (2015). Cross-language asymmetries in code-switching patterns: Implications for bilingual language production. *The Cambridge Handbook of Bilingual Processing*, 416–458.
- Sime, D. A. (2019). Code-switching in Ethiopian primary and secondary EFL classrooms: A comparison of its extent and types. *Journal of Foreign Language Education and Technology*, 4(2), 242–268.
- Temesgen, A., & Hailu, E. (2022). Teachers' codeswitching in EFL classrooms: Functions and motivations. *Cogent Education*, 9(1), 2124039.
- Wei, L. (2009). *Code-switching and the bilingual mental lexicon* in Bullock, B. E., & Toribio, A. J. E. (2009). *The Cambridge handbook of linguistic code-switching*. Cambridge University Press.

Acknowledgment

We sincerely thank Prof. Judit Navracsecs for proofreading, labeling examples, and editing our article.

ANITA KISS

University of Nyíregyháza
Department of Hungarian Linguistics
Institute of Languages and Literature
anita.kiss@nye.hu
<https://orcid.org/0000-0002-5087-3294>

Anita Kiss: Contrastive analysis of attitudes towards dialects
Alkalmazott Nyelvtudomány, XXV. évfolyam, 2025/1. szám, 122–136.
doi:<http://dx.doi.org/10.18460/ANY.2025.1.007>

Contrastive analysis of attitudes towards dialects

In Hungary and Transcarpathia, many studies have been carried out about the attitudes towards dialects among university students (cf. Kiss, 2009; Parapatics, 2021; T. Károlyi, 2002; Zimányi, 2015). The results of these examinations show a vastly different picture of the judgement of dialects. Some research displays an accepting attitude, while others display a rather dismissive attitude. In recent years, I have examined the attitudes of first-year teacher students at the University of Nyíregyháza towards dialect phenomena, which revealed that respondents do not hold negative opinions about dialects and are aware of the importance of conscious language use (cf. Kiss & Somfalvi, 2024; Kiss, 2025). This paper presents a comparative study of the dialect attitudes of Hungarian and Transcarpathian Hungarian students. The survey results showed that there were no significant differences in the dialect attitudes of the students of the two groups. This result can be explained by the fact that the dialects of both Transcarpathia and Szabolcs-Szatmár-Bereg county (where the majority of the Hungarian respondents live) belong to the North-Eastern dialect region, and thus show similar linguistic characteristics, and by the fact that the acceptance and respect of different language varieties is not only a local phenomenon, but also a wider one, regardless of geographical origin. The results of the study also showed that linguistic attitudes are not only a consequence of social factors, but also part of the linguistic culture, which needs to be consciously developed.

Keywords: language use, dialects, language attitudes, language diversity, language version

1. Introduction

Examining the speakers' attitudes and awareness connected to language variety and revealing their knowledge about conscious language use is a significant part of linguistic research. It is highly intriguing to observe this question among Hungarians living both within and outside Hungary's borders, as they live in different languages and cultural environments. Therefore, the social, economic, political, and cultural circumstances have an impact on their attitudes towards their language use, as well as the use of different languages and dialects.

Examining language attitudes is not only a linguistic matter; it is also connected to the formation of language identity. The identity-forming role of languages is especially pronounced in societies where more languages and dialects are used. Speakers with their language-related decisions and attitudes not only reflect on their own community, but they also express their relations and define their language identity with different dialects.

This study examines consciousness in language use and attitudes towards dialects among Hungarian students studying in tertiary education in Hungary and Transcarpathia. The purpose of this research is to reveal the differences between the participants of the two groups, considering their attitudes towards language diversity and dialects, as well as their thoughts on conscious language use.

The results of the research may help widen the literature on language attitudes and language diversity. Moreover, it may provide insight into the identity-forming processes of communities with different language backgrounds. The topic is relevant, since these days, due to increasing mobility, members of mono- and bilingual Hungarian communities can easily connect, and their dialects can have an effect on each other. We can achieve significant results from the connections between Hungarians living in Hungary and Transcarpathia, as well as from the language attitudes of students studying in higher educational institutions, which play an essential role in forming language consciousness.

2. The definition of language attitude

The attitude is a mental state of readiness, which is organized through experience and strongly influences the individual's reaction and behaviour (Allport, 1935: 810). In the conceptualisation of social psychology, attitude is a cognitive representational process that summarises the individual's evaluations of a particular person, group, thing or action (Domonkosi, 2004: 25). In general, attitude refers to beliefs and behaviours towards people, situations and ideas (Cseresnyési, 2004: 124; Fishbein & Ajzen, 1975: 6).

The knowledge and beliefs associated with language are collectively referred to as language attitude. Different studies approach the concept of language attitude from various perspectives. Some research have defined language attitudes as a set of evaluative judgements about a language or a language variety and, more particularly, about its speakers (Ryan et al., 1982: 7). Other studies do not only include value judgements about entire language varieties, but also about different linguistic phenomena such as the judgement of code-switching (cf. Dewaele & Wei, 2014). Thus, language attitude is the behaviour or a judgement of a person or a community towards a language, a variety of languages, a variety of pronunciations, or any linguistic phenomena which can be positive or negative (Grin, 2013: 682; Trudgill, 1997: 58).

Language attitudes originate in societies, not in languages; consequently, they express social habits, behavioural rules, and prejudice instead of linguistic or aesthetic values (Kiss, 1995: 136). They are not only individual but also social in nature; thus, they can express group identity, belonging to a community, or even separation from the group (Domonkosi, 2004: 25–26). Peter Trudgill's (1974) research on language use in Norwich reveals that, in addition to the high social prestige of the standard variety, the examined communities also attach prestige to

their own non-standard varieties of the language and are attached to their own variety, even if it has been negatively discriminated against.

Different language attitudes can influence the changes in the language. Negative attitudes towards certain linguistic features, such as dialectal phenomena, can result in insecurity in the speaker's language use. This can cause speakers to feel antipathy towards their own dialect and to question its correctness. As a consequence, they may try to acquire a higher-status form of speech (Trudgill, 1997: 58). According to Klára Sándor (2016: 96), it is not usually the language form that we have a negative attitude towards, but the groups within which it is used. Thus, our attitudes towards language do not actually refer to the language itself, but to the community or individuals behind it. This can also be directed towards the individual's own language variety. If someone is ashamed of using a language due to some negative grievance, they will try to distance themselves from that community and adapt their speech as much as possible to the standard. Linguistic attitudes are not static, but rather dynamic phenomena, since if the status of an individual or community changes, attitudes may also change (cf. Dewaele, 2009; Kircher & Zipp, 2022: 6).

Therefore, language attitude research is a crucial part of linguistic research. This paper focuses on the study of language awareness and language attitudes among students in higher education because, as future teachers, they can play a significant role in reducing the language insecurity of students who speak dialects and teaching them to use their language consciously.

3. Research about attitudes towards dialects in Transcarpathia and Hungary

There are different varieties of languages, and these varieties are what make a language diverse. The varieties of Hungarian are not entirely separate from each other, but are connected in the way people use the language, forming a network of languages. Each of them has its own role, usage, and linguistic specificity. The most apparent differences are mainly in their spread, function, social prestige, and vocabulary. Dialects play a crucial role in the diversity of language due to their variety (Kiss, 2017: 199). Hungarian dialects are language varieties that differ from the Hungarian standard at all levels of the language and are bound to a specific regional area (Beregszászi & Dudics Lakatos, 2023: 86–87). Most Hungarians still use dialects today, but the convergence to the vernacular is faster among Hungarians in Hungary than among minority Hungarians in the Carpathian Basin (Kiss, 1995: 194; Csernicskó & Szabó Mihály, 2023: 211). Language is changing constantly as a result of social, economic, technological, and cultural developments, and this also affects dialects. The change in dialects is reflected in the narrowing of their scope of use and a reduction in their regional specificity. This can be explained by the increasing need or demand for mobility, including intra-country and rural-to-urban population movements, as well as commuting.

Therefore, dialect speakers are increasingly trying to acquire a standard language, as a consequence of which, depending on the situation, they use the standard or a regional version of it in communication, which is referred to as a regional standard (Kiss, 2017: 200–201).

Research has shown that the individual's attitude towards their language use can be influenced by others' beliefs about that particular language variant. For instance, if speakers are judged or despised because of their language use, they can form a negative impression not only of their dialect, but also of their own community (Sándor, 1999: 165–168). Therefore, over the last few decades, a pluricentric approach, which accepts language variations, has become increasingly important. It means that pursuing additive, functional-situational bilingualism could change dialectal speakers into more conscious language users. It could alter negative preconceptions (e.g., stigmatization) connected to dialects (Parapatics, 2022: 85). The terms monocentric and polycentric, used to describe national multilingualism, were introduced by William Stewart (1968). These two concepts represent different ways of describing the process of standardisation. The monocentric view is that there is a common standard, which does not take into account linguistic diversity. In the case of pluricentric languages, there is no single central norm that determines the 'correct' use of the language; rather, several national or regional varieties coexist and are of equal status (cf. Clyne, 1992).

There are differences in language use between Hungarians and Transcarpathian-Hungarians. One of these typical alterations is that the use of the Hungarian language in Transcarpathia contains more dialectal elements than in the motherland. Moreover, there are additional regional components that deviate from the standard, and the archaic forms persist for a more extended period. Additionally, members of the minority Transcarpathian-Hungarian community use dialects in situations where members of the motherland no longer use them anymore (Kiss, 1995: 192–194).

The examination of the Transcarpathian language use from a sociolinguistic approach started in the second half of the 1990s. Questions about language attitudes and the judgement of language variants are present in a series of research that involves the whole Carpathian Basin. The examination of the language use of Hungarians living in Transcarpathia shows that the Transcarpathian-Hungarian community has a positive attitude towards their mother tongue. For them, the Hungarian language serves as an identity marker, which holds significant importance, while the local language (Transcarpathian-Hungarian) expresses regional attachment (cf. Csernicskó, 1998).

Two surveys have been conducted to assess attitudes towards dialects among pedagogues. According to the results of the former survey, teachers acknowledged that people in their surroundings speak in dialects; however, they exhibited a distant attitude connected to dialectal speech (Lakatos, 2010: 146–

170). Ten years later, they conducted the survey again. By then, the promotion of a paradigm shift in mother tongue education (pluricentric approach) had happened. According to the results of this survey, a positive change had begun by that time, although a separating attitude towards dialectal language use was still typical (cf. Dudics Lakatos, 2019).

A study examining language consciousness and attitudes among higher education students in Transcarpathia revealed that the examined teacher-training students are aware of the pluricentric approach, pay attention to situational language use, and consistently use the standard forms in proper situations (T. Károlyi, 2002: 329–333). Another study revealed that the language-related value judgement of university students is influenced by having a linguistic education during their higher education studies. Moreover, the participants' positive attitude towards the use of the Transcarpathian-Hungarian language was perceptible. Most participants find their own settlement's language exquisite, but they are also aware that it differs from the motherland's version of the Hungarian language (Dudics Lakatos & Gazdag, 2023: 39–41).

In Hungary, the language-related value judgement was examined in the Hungarian population in the framework of the Hungarian National Sociolinguistic Research in 1988. According to the results, individuals with higher qualifications are more likely to consider a non-standard sentence grammatically incorrect. Participants living in the capital city marked the non-standard sentences correct in a smaller proportion than those who live in the countryside. Moreover, it was also revealed that women evaluated standard sentences as correct forms in more cases than men (Kontra, 2003: 577–594).

Surveys about attitudes were also conducted among university students. According to a 2009 research study, 92% of university students in Budapest do not have a dialectal background (Kiss, 2009: 3). In a later survey, attitudes towards dialects were examined among university students in Eger and Budapest. It revealed that two-thirds of Hungarian faculty students in Eger, and almost half of the students in Budapest, acknowledged that they had never had direct contact with any dialects, while the students from other faculties answered yes to the same question. It was also shown that most students in Eger accept dialects and find them equal to the standard form. Among the students in Budapest, a larger percentage held a negative belief about the dialectal phenomena (Zimányi, 2015: 234–237).

According to the results of the language-related attitude research conducted at the University of Debrecen, students in the pedagogy faculty consider the standard language appropriate in the school environment, although they also accept teachers using dialects in their lessons (cf. Kovács, 2014).

Andrea Prapatics (2021) examined students' awareness of dialects and language at various Hungarian universities. She analysed her data based on the national pattern and the answers of participants at Pannon University. She concluded that

the mentality of university students is standard-centered; they are less accepting of regional language forms, the reason for which is probably the lack of knowledge and consciousness connected to their mother tongue.

The results above show that both the Hungarian and Transcarpathian-Hungarian students' opinions about the judgement of territorial variety of language use show a varied picture. Indeed, students also have the experience that dialects are often subject to negative judgements, due to misconceptions, which affects their attitudes towards these dialects. Consequently, it is worthwhile to examine the question of language consciousness and the use of dialects in secondary schools and higher education as well.

4. The methods and participants of the research

This research is a comparative analysis involving university students from Transcarpathia and Hungary. The aim of this study is to examine the differences that occur between the two groups in terms of language attitude and how they relate to language diversity and dialects.

The research was conducted through an online survey completed by 108 students in teacher training. Questionnaire data were collected in 2024. I have previously undertaken a dialect attitudes survey among students at the University of Nyíregyháza (cf. Kiss & Somfalvi 2024; Kiss, 2025). This study presents a new research study involving university students who have not previously participated in the survey. In addition, I would have liked to extend the study to students studying in other higher education institutions, and therefore, Hungarian students from Transcarpathia were also involved in the research. The members of the Transcarpathian group are studying at Ferenc Rákóczi II. The Transcarpathian Hungarian College of Higher Education, while the Hungarian participants are students of the University of Nyíregyháza. The reason for choosing students from teacher-training programmes was that, as future teachers, they can help decrease the uncertainty of language use among students who use dialects; moreover, they will also be able to teach them conscious language use with an academic purpose.

The questionnaire was based on the results of similar studies published in the literature (Lakatos, 2010; Kiss, 2009) and on my previous experiences of collecting data. The questions were divided into two blocks. The first part concerned the personal background information of the respondents (gender, age, place of birth, current place of residence). The second part of the questionnaire was about dialect attitudes. The closed-ended questions were intended to find out whether respondents notice dialect features in their own and others' language use, how they relate to dialect features, in what situations they use them, and in what situations they think dialect features can be used. The open question asked respondents to define the term 'dialect'. This question was important in part to find out how students understand the term. On the other hand, this question also allowed us to gauge the respondents' subjective opinions about dialect speech.

The questionnaire also included Likert scale questions to examine their knowledge of conscious language use and their attitudes towards dialect. The questionnaire primarily consisted of closed-ended questions to ensure data comparability.

5. Research hypotheses

For the plans of the research, based on the results of previous empirical examinations, I listed the following hypotheses:

1. Research about the language use of Hungarians living in Transcarpathia has shown that for them, the identity marker function of the language is more important than for those Hungarians who live in the motherland (cf. Csernicskó, 2008). Therefore, I presume that the Transcarpathian-Hungarian participants will have a more positive attitude towards language diversity and dialects than the Hungarian respondents.

2. Earlier research's results about attitudes towards dialects show an immensely varied picture (cf. Dudics Lakatos, 2019). Both Hungarian and Transcarpathian-Hungarian studies partly report on participants' favorable value judgement connected to dialects, while others discuss less accepting attitudes (cf. Kiss, 2009; Parapatics, 2021; T. Károlyi, 2002; Zimányi, 2015). According to these, it is assumable that the Hungarian and Transcarpathian participants of my research group will also have different beliefs about dialects.

3. Studies connected to dialects have revealed that in most cases, participants are unaware of using dialects in their speech (cf. Dudics Lakatos, 2019; Sándor, 2015). I presume that among students in the research group, I will observe insecurities in distinguishing between standard language variants and dialectical forms.

In the research, I also wanted to find answers to the following questions:

1. How do the participants define dialects?
2. In which situations do they use these dialects?
3. How do they judge dialects in language use?
4. Do they recognize standard and dialectical language forms?
5. What type of knowledge do they have about conscious language use?
6. What are the differences in the dialect attitudes of Hungarian and Transcarpathian respondents?

6. Results

According to previous results, Hungarians living in Transcarpathia have a favorable opinion about their dialects. In several studies, participants were asked to choose the region where they believed the most exquisite version of Hungarian is spoken. In a 1996 study, 55,9% of the participants chose Transcarpathia, and 39% of them picked Hungary (cf. Csernicskó, 1998). In a following survey, two-thirds of the participants also marked Transcarpathia (Karmacsi, 2009: 415–422).

Similar results were shown in another survey, involving students from Transcarpathian higher educational institutions, in which 55,2% of the participants claimed that the most exquisite version of Hungarian is spoken in Transcarpathia (Gazdag & Dudics Lakatos, 2023: 39–40). In my own research, the participants were requested to answer the same question. The answers are shown in the first table. Similarly to previous results, the majority of participants (56%) identified Transcarpathia, while 38% of them were unable to judge it. The majority of Hungarian participants (44%) believe that the most beautiful version of the language is spoken in the Hungarian countryside, while 33% of them were unable to express an opinion about it. According to the Hungarian National Sociolinguistic Research in 1988, the majority of participants believed that the most exquisite version of Hungarian is spoken in Budapest. On the other hand, participants from the north-east language area thought that their own dialect is the most beautiful (Kontra, 2003: 244). In my research, the majority of participants come from the north-east language area, and the results show that only 15% of them believe that the best version of Hungarian is spoken in Budapest.

Table 1. Respondents' opinions on where people speak Hungarian best

Where do you think the most beautiful version of Hungarian is spoken?	Transcarpathians		Hungarians	
	pc	%	pc	%
I couldn't judge	21	38	18	33
in Budapest	1	2	8	15
in the Hungarian country towns	1	2	23	44
in Hungarian villages	1	2	4	7
in Transcarpathia	30	56	1	1
Total:	54	100	54	100

After this, I asked the participants an open-ended question to determine whether they were aware of the definition of dialect. In attitude studies, it is essential to ensure that participants can recognize the subject of the attitude. All participants answered the question, and they mostly mentioned the dialect's variety, the geographical boundaries, dialectical idioms, and the differences in pronunciation. The participants tried to formulate their answers based on metalinguistic knowledge, which in none of the cases contained obviously positive or negative assumptions. For example: 'The specific vocabulary of a given area is called dialect, which is an important part of the personality', 'Dialect is the variety of language that differs from the standard in pronunciation and vocabulary', 'People speak differently from one region to another, use different regional words, speak with different pronunciation', 'A dialect is a variety of a language spoken in a particular area or community. These variants may differ in pronunciation,

vocabulary, and grammatical rules'. A small percentage of the participants answered that mostly elderly people living in villages speak in dialects: 'The dialects contain words that are mainly used in villages. They differ from the standard', 'People speak and pronounce words differently in various communities. In some places, people speak the old way, especially the elderly'. People from different geographical areas speak diverse forms of the same language. These differences have already been studied by linguists, who have come to the conclusion that speakers are aware of these variants. They may also be able to recognize the area from which the speaker originates (cf. Clopper & Pisoni, 2007; Clopper et al., 2012; Wagner et al., 2013).

In the survey, respondents were asked if they noticed any differences between their language use and the neighbouring settlement's language use. The majority of the answers were yes in both groups; however, the Hungarian participants (80%) claimed it at a higher percentage than the Transcarpathian-Hungarian group (65%). Comparing the responses in Hungary with the results of first-year students in the academic year 2023 (Kiss, 2025: 85), it is evident that the present study found an increase in the number of students who answered yes to the question.

The answers revealed that Hungarian (72%) and Transcarpathian (62%) participants believe that dialects are also used in their own settlements. Furthermore, the majority of students acknowledged that dialectical forms are present in their own language use as well, as 52% of Hungarians and 61% of Transcarpathians answered yes to this question. It is interesting to compare the views of the Hungarian respondents with the results of the 2023 sample. In the previous survey, only 39% of respondents said yes, whereas in the current survey, more than half of the respondents shared the same opinion. The responses were analysed using a T-test, which showed a significant difference between the two groups' opinions: $t(117)=-3,045$ $p>0,05$ ($p=0,002$). This result indicates that the Hungarian respondents in the present study recognised the occurrence of dialect phenomena in their own language use to a much greater percentage than the Hungarian group previously studied (Kiss, 2025: 88).

Participants mostly use their dialects in friend or family groups (Transcarpathians: 43%, Hungarians: 31%), which proves that they make an effort to use the language more consciously. Some participants were not able to decide in which situations they use the dialectical forms (Transcarpathians: 24%, Hungarians: 28%). Moreover, only a small percentage of the participants acknowledged that they use their dialects everywhere (Transcarpathians: 11%, Hungarians: 7%). Participants in Transcarpathia (67%) and also in Hungary (61%) mostly believe that it is acceptable to use dialects in every situation (Table 2).

Table 2. Respondents' views when someone speaks in dialect everywhere

What do you believe about people who use dialects always and everywhere?	I find it right	I find it acceptable	I do not find it right	Total
Transcarpathians	12 (22%)	36 (67%)	6 (11%)	54 (100%)
Hungarians	9 (17%)	33 (61%)	12 (22%)	54 (100%)

On the other hand, participants thought more positively about those who pay attention to situational language use. It means that participants do not find it wrong when people use dialectical forms with their friends and family members, while they speak in standard forms in formal situations (Table 3).

Table 3. What is your opinion about people who use their most natural language forms (e.g., dialects) among friends and family members, while they speak in standard forms in formal situations?

	I find it right	I find it acceptable	I do not find it right	Total
Transcarpathians	27 (50%)	27 (50%)	0%	54 (100%)
Hungarians	36 (67%)	15 (28%)	3 (5%)	54 (100%)

The research revealed that 57% of Transcarpathian and 61% of Hungarian students consider dialects and standard forms equal. This result is similar to the opinion of students at Eszterházy Károly Catholic University, the majority of whom also find standard and dialectical forms equal (Zimányi, 2015: 237).

The two groups of participants had similar views about the definition of dialects. The Hungarian respondents mostly (59%) believe that dialects are used to express identity and communication. The majority of Transcarpathian students (72%) also emphasised the function of expressing identity, and 59% of them indicated that dialects are a tool of communication. However, there were more participants in both groups (Hungarian students: 35%, Transcarpathian-Hungarian students: 41%), who believe that mostly those who use dialects live in villages and work in agriculture. Moreover, some of them (Hungarian students: 26%, Transcarpathian-Hungarian students: 31%) also marked that dialectical language use is more typical among the elderly. Only a few participants had an explicitly negative opinion about dialects.

One purpose of the examination was to determine whether participants could distinguish between standard and dialectical vocabulary. In the questionnaire, I set a list of words that contained both standard and dialectical forms. Students were requested to mark those words that they considered dialectical. The dialect words are in bold in Table 4. Overall, both groups recognized the majority of dialectical words, although some mistakes were made by both the Hungarians and Transcarpathians as well.

In some cases, they identified standard words dialectal forms, for instance *krumpli* ‘potato’, *tied* ‘yours’, *dobjál* ‘throw in second person imperative form’, and *leány* ‘girl’, which have two variants in standard speech. It also occurred that they marked a dialectal form standard, for instance, the word *pocok* (vole). The answers show that only a few of the Transcarpathian Hungarian respondents considered the form *kellesz* to be dialectal, while the majority of the Hungarian respondents correctly recognized the dialect word. Also in the previous survey, the majority of Hungarian respondents correctly recognised the dialect word *kellesz* (Kiss 2025: 88). This could indicate that among Transcarpathian Hungarians, the word *kellesz* is so common in the community's language use that they do not feel its dialectal character. Similarly, the dialect words *jösztek*, *lábtó*, and *tallu* were correctly classified as dialect words by Hungarians at higher percentages than by Transcarpathians. However, the standard form *tied* was incorrectly classified by more Hungarians than by Transcarpathians.

Table 4. Standard and dialectal words

Which words do you think are dialectal? Mark them.	Transcarpathians%	Hungarians%
<i>furik</i> 'wheelbarrow'	87	94
<i>jösztek</i> 'come in second person plural'	61	87
<i>kokas</i> 'cock'	70	87
<i>lábtó</i> 'ladder'	54	80
<i>tallu</i> 'feather'	56	78
<i>kellesz</i> 'must'	31	74
<i>tied</i> 'yours'	41	63
<i>leány</i> 'girl'	31	48
<i>dobjál</i> 'throw'	39	37
<i>krumpli</i> 'potato'	59	30
<i>pocok</i> 'vole'	39	28
<i>gyermek</i> 'child'	6	6
<i>kánikula</i> 'heatwave'	30	6
<i>örökség</i> 'heritage'	2	0

In the following questions, participants were asked to mark on a 1–5 scale the extent to which they agreed with various statements. In the questionnaire, only two extreme rates were provided: 1 = "I do not agree at all" and 5 = "I absolutely agree."

The majority of Transcarpathian students (61%) marked 3, indicating a neutral opinion about the statement that these days only a few people speak in dialects. The Hungarian students mostly marked three as well (43%), although they were followed closely by those (33%) who agreed less with the same statement.

According to the second statement, educators are responsible for teaching students the standard forms, rather than dialects. In this statement, the emphasis is on “instead of”, which refers to the method of mother-tongue education, following a one-standard, replacing approach. The Transcarpathian participants marked 3 (32%) and 4 (33%) almost equally, indicating that some of them had held a neutral opinion about this statement as well, while the other part instead agreed with it. The majority of Hungarian participants (41%) instead agreed with the statement. The following statement was formulated similarly to the previous one: “The educators are responsible for drawing students’ attention to the value of dialects besides the standard forms.” In this statement, the word “beside” is emphasized, which refers to the pluricentric approach. Most of the Transcarpathian students (39%) and the majority of Hungarian respondents (54%) strongly agree with this statement. Consequently, students involved in the research mostly prefer the additive mother-tongue educational approach. Transcarpathian students held various opinions regarding the statement that, in these days, acquiring standard forms is extremely important. 35% of the respondents held a neutral view, 30% absolutely agreed with it, and 22% of the participants somewhat approved of it. The Hungarian respondents' opinions were more united, as the majority of them (52%) completely agreed with the statement. According to the following statement, people who speak in dialects are uneducated. Half of the Transcarpathian students disagreed with this statement, and 70% of Hungarians shared the same opinion. In both groups, the majority of respondents (Transcarpathian students: 37%, Hungarian students: 35%) completely agreed that it is beneficial to speak in one of the dialects and use standard forms as well. Moreover, both Transcarpathian-Hungarian (52%) and Hungarian (56%) students mostly agreed that dialects are valuable and it is crucial to preserve them.

7. Summary

This research focused on the dialect attitudes and language consciousness of Hungarian university students in Hungary and Transcarpathia. As a starting point for the study, it was presumed that Transcarpathian-Hungarian participants would have a more positive attitude towards language diversity and dialectal language use than the Hungarian respondents (1st hypothesis). The results did not support this hypothesis, as the respondents of the two groups held similar beliefs about language diversity, and there were no significant differences in their value judgements. This can probably be explained by the fact that both Transcarpathia and Szabolcs-Szatmár-Bereg county (where the respondents’ majority live) belong to the north-east language region, therefore they show similar dialectal characteristics.

In general, both Transcarpathian-Hungarian and Hungarian students had a favorable opinion about dialectical language use. Answers referring to negative attitudes could not be found in either group (2nd hypothesis). In both groups, the participants believe that dialectical forms are typical in their own settlement; moreover, they also acknowledge that dialects are present in their own language use. In most cases, the participants could distinguish between standard and dialectical vocabulary; they only identified a few words incorrectly, most of which have two standard variants. The respondents' knowledge of dialects is also confirmed by the fact that in both groups, the answers written by the students mostly covered the definition of the term as described in the literature summary. These results do not support hypothesis 3. Participants think it is essential to strive to use language consciously. The majority of respondents admit to speaking in dialect only among family and friends. However, they are not rejecting dialect speech, and agree that students should be taught about language varieties in an additive way at school. Hopefully, the participants involved in this research will be able to apply the appropriate approach as pedagogues, teaching students about language diversity and fostering a more positive value judgement in students regarding dialects and language diversity.

On the whole, the research broadens dialectical attitude examinations with important results; furthermore, it can also contribute to the understanding and acceptance of conscious language use and language diversity. The comparison of the attitude towards dialects in these two culturally and geographically dissimilar groups (Transcarpathian-Hungarian and Hungarian students), and the results received also show that language diversity and the openness towards dialects do not only depend on regional identity, but they are also based on the knowledge and experiences obtained in secondary and higher education.

The conclusion is that there were no significant differences in the attitudes of the two groups towards dialects. It also reveals that accepting and respecting different dialects is not only a regional phenomenon, but also a widely characteristic trait, regardless of geographical affiliation.

A significant aspect of the research is that attitudes towards dialects are not only a consequence of social circumstances, but they are also an integral part of the culture, which needs to be improved consciously. The favorable judgement of dialects and conscious language use is critical in the preservation and maintenance of languages. These are not only crucial at the level of language, but also in strengthening social relationships and cultural identity.

I plan to continue and expand the research in the future. Firstly, I would like to involve teachers in the study and examine their attitudes and knowledge of dialects, as well as the approach they take to teaching students about linguistic diversity. On the other hand, since this questionnaire is completed annually by first-year teacher students at the University of Nyíregyháza, I also plan to conduct a study comparing the results obtained from year to year. This may reveal whether

attitudes towards linguistic diversity and perceptions of dialects are changing positively or negatively.

Acknowledgements

Supported by the EKÖP-24 University Excellence Scholarship Program Of The Ministry For Culture And Innovation From The Source Of The National Research, Development And Innovation Fund. Supporting institution: University of Nyíregyháza.

References

- Allport, G. W. (1935). Attitudes. In: Murchison, C. (ed.), *A handbook of social psychology*. Worcester, MA: Clark University Press. 798–844.
- Beregszászi, A. & Dudics Lakatos, K. (2023). Huszonkettő - A kárpátaljai magyar anyanyelvi nevelés 22 évének története. Törökbálint: Termini Association.
- Clopper, C. G. & Pisoni, D. B. (2007). Free classification of regional dialects of American English. *Journal of Phonetics* 35. 421–38. <https://doi.org/10.1016/j.wocn.2006.06.001>
- Clopper, C. G. & Rohrbeck, K. L. & Wagner, L. (2012). Perception of dialect variation by young adults with High-Functioning Autism. *Journal of Autism and Developmental Disorders* 42. 740–54. <https://doi.org/10.1007/s10803-011-1305-y>
- Clyne, M. (1992). Introduction. In: Clyne, M. (ed.), *Pluricentric Languages. Different Norms in Different Countries*. Berlin/New York: Mouton de Gruyter. 1–9.
- Cseresnyési L. (2004). *Nyelvek és stratégiák avagy a nyelv antropológiája*. Budapest: Tinta Könyvkiadó.
- Csernicskó, I. (1998). *A magyar nyelv Ukrajnában (Kárpátalján)*. Budapest: Osiris–MTA Kisebbségkutató Műhely.
- Csernicskó, I. (2008). Nyelv és azonosságtudat összefüggései a kárpátaljai magyar közösségben. In: Fedinec, Cs. (szerk.), *Értékek, dimenziók a magyarságkutatásban*. Budapest: Magyar Tudományos Akadémia Magyar Tudományosság Külföldön Elnöki Bizottság. 153–170.
- Csernicskó, I. & Szabó Mihály, G. (2023). Hátrányból előnyt: a magyar nyelvpolitika és nyelvtervezés kihívásairól. In: Prószték, G. (főszerk.); Fóris, Á. & B. Papp, E. & Bölskei, A. & Lipp, Veronika (szerk.), *A magyar terminológiastratégia kialakítása: Zöld könyv*. Budapest: Nyelvtudományi Kutatóközpont. 199–235.
- Dewaele, J. M. (2009). Perception, attitude and motivation. In: Cook, V. & Wei, L., (eds.), *Language Teaching and Learning*. London: Continuum, 163–192.
- Dewaele J. & M., Wei L. (2014). Attitudes towards code-switching among adult mono- and multilingual language users. *Journal of Multilingual and Multicultural Development* 35/3: 235–251. <https://doi.org/10.1080/01434632.2013.859687>
- Domonkosi Á. (2004). A nyelvi attitűdök és a nyelvhasználat szabályozottsága. In: Kurtán Zs. & Zimányi Á. (szerk.), *A nyelvek vonzásában. Köszöntő kötet Budai László 70. születésnapjára*. Eger–Veszprém: Veszprémi Egyetemi Kiadó. 25–31.
- Dudics Lakatos, K. (2019): Kárpátaljai magyar pedagógusok nyelvjárási attitűdjének alakulása két felmérés eredményei alapján (2008–2018). *Magyar Nyelvjárások* 57. 123–134. DOI: 10.30790/mnyj/2019/08
- Dudics Lakatos, K. & Gazdag, V. (2023). A kárpátaljai felsőoktatási intézmények hallgatóinak nyelvhasználati sajátosságai és nyelvi attitűdjei két felmérés részeredményei alapján. *ACTA Academiae Beregsasiensis. Philologica* II. 32–48. <http://dx.doi.org/10.58423/2786-6726/2023-2-32-48>
- Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behaviour. An introduction to theory and research*. Reading, MA: Addison-Wesley [ebből magyarul olvasható: 222–244, 253–255, Az attitűd információs alapjai. In: Hunyady, Gy. (szerk.), 1984. *Szociálpszichológia*. Budapest: Gondolat. 175–210].

- Grin, F. (2013). Language Policy, Ideology, and Attitudes. In: Bayley, R. & Richard C. & Ceil L. (eds.), *The Oxford Handbook of Sociolinguistics*. Oxford: University Press. 629–650. <https://doi.org/10.1093/oxfordhb/9780199744084.013.0031>
- Karmacs, Z. (2009). Nyelvi attitűd: gyengítő és erősítő tényezők Kárpátalján a magyar nyelv esetében. In: Borbély, A. & Vančóné Kremmer, I. & Hattyár, H. (szerk.), *Nyelvideológiák, attitűdök és sztereotípiák*. Budapest–Dunaszerdahely–Nyitra: MTA Nyelvtudományi Intézet–Gramma Nyelvi Iroda–Konstantin Filozófus Egyetem Közép-európai Tanulmányok Kara. 415–422.
- Kircher, R. & Zipp, L. (2022). An introduction to language attitudes research. In: Kircher R. & Zipp, L. (eds.), *Research Methods in Language Attitudes*. Cambridge: University Press. 1–16.
- Kiss, A. (2025). Nyelvjárási tudat és attitűd az egyetemi hallgatók körében. *Magyar Nyelvőr* 149/1: 80–93.
- Kiss, A. & Somfalvi, Z. (2024). Language awareness survey among university students. *Acta Academiae Beregsasiensis. Philologica* 3/2: 53–69.
- Kiss, J. (1995). *Társadalom és nyelvhasználat*. Budapest: Nemzeti Tankönyvkiadó.
- Kiss, J. (2009). A nyelvjárások és a dialektológiaoktatás Kárpát-medencei magyar szakos hallgatók szemével. *Magyar Nyelvőr* 133. 1–14.
- Kiss, J. (2017). A nyelvjárások. In: Tolcsvai Nagy, G. (szerk.), *A magyar nyelv jelene és jövője*. Budapest: Gondolat Kiadó. 199–221.
- Kontra, M. (szerk.) (2003). *Nyelv és társadalom a rendszerváltás kori Magyarországon*. Budapest: Osiris Kiadó.
- Kovács, E. (2014). A tanár szakos hallgatók nyelvi attitűdjének társadalmi nemi sajátosságai. *Anyanyelv-pedagógia* 7/2. <https://www.anyanyelv-pedagogia.hu/cikkek.php?id=514>.
- Lakatos, K. (2010). Kárpátaljai magyar iskolások nyelvi tudata és attitűdje. Kézirat. *Doktori értekezés*. Budapest: ELTE BTK.
- Parapatics, A. (2021). *Pannon Egyetemen tanuló hallgatók nyelvi-nyelvjárási tudata*. In: *Humán tudományok: pedagógusképzés és tananyagfejlesztés*. Budapest: Akadémiai Kiadó.
- Parapatics, A. (2022). Nyugat-magyarországi beszélők nyelvi-nyelvjárási attitűdjei. *Részeredmények. Acta Academiae Beregsasiensis. Philologica* 1. 85–105. <http://dx.doi.org/10.58423/2786-6726/2022-1-85-105>
- Ryan, E. B. & Howard G. & Sebastian, R. J. (1982). An integrative perspective for the study of attitudes toward language variation. In: Ryan, E. B. & Howard G. (eds.), *Attitudes towards language variation*. London: Edward Arnold. 1–19.
- Sándor, K. (1999). Szociolingvisztikai alapismeretek. In: Galgóczi, L. (szerk.), *Nyelvtan, nyelvhasználat, kommunikáció*. Szeged: JGYF Kiadó. 133–171.
- Sándor, K. (2015). Hogy keletkeztek a "nagyon nyelvjárások"? *Nyelv és Tudomány* 2015. május 19. <https://www.nyest.hu/hirek/hogy-keletkeztek-a-nagyon-nyelvjarasok> Letöltve: 2025. 01.07.
- Sándor K. (2016). *Nyelv és társadalom*. Budapest: Krónika Nova.
- Stewart, W. (1968). A sociolinguistic typology for describing national multilingualism. In: Fishman, J. A. (ed.), *Readings in the sociology of language*. The Hague: Mouton de Gruyter. 530–545.
- T. Károlyi, M. (2002). A nyelvi tudatosság és attitűd vizsgálata a beregszászi főiskola hallgatóinak különböző csoportjaiban. In: Hoffmann, I. & Juhász, D. & Péntek, J. (szerk.), *Hungarológia és dimenzionális nyelvszemlélet*. Debrecen: Debreceni Egyetem Magyar Nyelvtudományi Tanszék. 329–338.
- Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Trudgill, P. (1997). *Bevezetés a nyelv és társadalom tanulmányozásába*. Szeged, JGYTF.
- Zimányi, Á. (2015). A nyelvjárások és a nyelvjárásiaság megítélése egyetemisták és főiskolások körében. *Acta Academiae Paedagogicae Agriensis Nova Series: Sectio Linguistica Hungarica* 42. 233–244.
- Wagner, L. & Clopper, C. G. & Pate, J. K. (2013). Children's perception of dialect variation. *Journal of Child Language* 41. 5: 1–23. <https://doi.org/10.1017/s0305000913000330>

CSILLA IBOLYA SÓLYOM

University of Debrecen, Doctoral School of Linguistics

e-mail: solyom.csilla@arts.unideb.hu

<https://orcid.org/0009-0002-3602-6133>

Csilla Ibolya Sólyom: Transfer of pronominal possessors, differential object marking, and resultative constructions in L3 acquisition of English

Alkalmazott Nyelvtudomány, XXV. évfolyam, 2025/1. szám, 137–160.

doi:<http://dx.doi.org/10.18460/ANY.2025.1.008>

Transfer of pronominal possessors, differential object marking, and resultative constructions in L3 acquisition of English

This paper presents the results of an experimental test conducted with two age- and proficiency-matched groups of English learners: an L2 group (L1 Hungarian L2 English; N = 20) and an L3 group (L1 Hungarian L2 Romanian L3 English; N = 24). The grammatical phenomena under investigation are pronominal possessors, differential object marking, and resultative constructions. The experimental test consists of an acceptability judgement task in English, designed to assess learners' comprehension of grammatical and ungrammatical target sentences in the initial stages of learning English as an L2/L3. The paper contributes to the field of L3 acquisition by the novel combination of languages involved in the experiment and the target grammatical conditions that have not been examined in this context before. The results provide evidence for property-by-property facilitative and non-facilitative linguistic transfer from both previously learned languages (predicted by the Linguistic Proximity Model by Westergaard et al. (2017) and the Scalpel Model by Slabakova (2017)).

Keywords: L3 acquisition, transfer, pronominal possessors, differential object marking, resultative constructions

1. Introduction

The main goal of this paper is to investigate the comprehension of pronominal possessors, differential object marking (DOM), and resultative constructions (RCs) in the initial stages (A2 level of proficiency) of L3 acquisition of English. An experimental test was conducted on two groups of English learners: an L2 group (L1 Hungarian, L2 English) and an L3 group (L1 Hungarian, L2 Romanian, L3 English). The purpose of comparing the two groups was to determine the source of linguistic transfer/crosslinguistic influence (CLI) – whether it be the L1, L2, or both¹. Hungarian (L1), Romanian (L2), and English (L3) comprise the language combination used in the investigation. The languages in question are typologically distant, falling under the Finno-Ugric, Romance, and Germanic language families.

¹ I use the terms transfer and CLI interchangeably, like other authors (e.g. De Angelis, 2007, p. 19; Odlin, 2003, p. 436).

The target grammatical properties with regard to the three languages involved in the study are discussed in Section 2.

As per the literature, the learner's L1 or L2 can serve as the source of transfer, which can either facilitate or impede the process of learning the intended L3. A variety of L3 transfer models have emerged in the last twenty years, each predicting that the source of transfer would be either the L1 or the L2, or both, of the formerly learned languages. Some models suggest that transfer can be beneficial as it facilitates the acquisition of L3. Others contend that transfer can also be detrimental (non-facilitative), thereby obstructing the acquisition of the target language.

The L1 scenario (Hermas, 2010, 2015; Na Ranong & Leung, 2009) anticipates that the L1 will be the favoured language for transfer. In contrast to the L1 scenario, the L2 Status Factor hypothesis (L2SF; Bardel & Falk, 2007; Falk & Bardel, 2011) posits that the second language is selected for transfer based on its psychological and cognitive significance, as well as its recent origin. L3 syntactic transfer from L1 or L2 is claimed by Cumulative Enhancement Model proponents (CEM; Flynn et al., 2004). Language acquisition is a sequential and incremental process, according to the CEM, and prior language knowledge may be beneficial or irrelevant to the learning process. The Typological Primacy Model (TPM; Rothman, 2011, 2013, 2015) anticipates a complete initial transfer from the language (L1 or L2) that is typologically closest. The model accommodates both positive and negative transfer. The parsing of the new input L3 grammar may require a longer or reduced duration, contingent upon the similarities between the languages.

Proponents of the Linguistic Proximity Model (LPM; Westergaard et al., 2017; Westergaard, 2020) argue that properties are acquired one at a time and permit the use of input from one or both previously acquired languages, allowing it to be both facilitative and non-facilitative. CLI arises when a linguistic feature of a language being taught is structurally analogous to features of languages that have been previously acquired. Nevertheless, this analogy may only be present at the initial stage of the learner's incorrect perception, which prevents the facilitation of a previous language. Similar to the LPM, the Scalpel Model (SM; Slabakova, 2017) predicts property-specific positive and negative transfer from the L1 and/or L2. Westergaard et al. (2017) and Westergaard (2020) emphasize the process of co-activation of all known languages in the learner's mind when acquiring an L3/Ln, whereas Slabakova (2017) emphasises the learner's capacity to make scalpel-like, sharp distinctions between language properties.

The paper aims to present how transfer occurs in the case of the two groups, to identify the source and driving factors of transfer, and ultimately provide support for or contradict predictions of the L3 acquisition models briefly discussed above in relation to the target grammatical properties and languages used in the experiment.

The paper is organized as follows: the target grammatical features in the three languages are introduced in Section 2. The experimental investigation is the primary focus of Section 3, while the results are presented in Section 4. Section 5 serves as the paper's conclusion.

2. Linguistic background

2.1. Pronominal possessors

The three languages manifest variations in the types of pronominal possessors that are the subject of the current study. One of the experiment's objectives was to investigate the possibility of employing the definite article with the third person singular possessive pronoun(s) *his/her* in English (as a result of linguistic transfer by the learners), as well as the variations in word order between Hungarian, Romanian, and English when expressing possession. Within the grammatical category of pronominal possessors, this paper focuses only on some specific forms of expressing possession (out of many), which can be regarded as equivalent and to a certain degree similar structures in the three languages: caseless possessor in Hungarian preceded by the definite article, genitive-marked personal pronouns in Romanian with suffixal definite article, and agreeing possessive pronouns in English which cannot occur with the definite article. The structures are briefly described and illustrated by examples in the following subsections.

2.1.1. Hungarian. The caseless possessor *ő* (his/her) is always preceded by the definite article *a/az* (the), as presented in example (1a). The caseless possessor *ő* (his/her) is optionally used in Hungarian (see example (1b)). However, its use is preferred when the aim is to express contrast.

- (1) a. *az ő könyve*
the he/she book.POSS
'his/her book'

Hungarian

(É. Kiss, 2002, p. 158, (14b))

- (1) b. *a könyve*
the book.POSS
'his/her book'

Hungarian

2.1.2. Romanian. According to Dobrovie-Sorin and Giurgea (2013), Romanian uses genitive-marked forms of the personal pronoun for the third person, which are gender-congruent with the possessor (i.e., a masculine or feminine head noun). The genitive case is employed to indicate third-person possessors, as illustrated in example (2).

- (2) fata lui / ei / lor Romanian
 girl.F.DEF he.GEN / she.GEN / they.GEN
 ‘his girl’ / ‘her girl’ / ‘their girl’

Another critical element in this context is the presence of the definite article in the form of a clitic affixed to the noun: *-a* for the feminine noun (example (2)) and *-le* for the masculine noun (as in *fratele lui* (his brother)). Additionally, it is important to note that the possessive pronoun is placed after the noun, which results in a word order distinction from English.

2.1.3. English. The inherent gender of the possessor is reflected in the third-person singular possessive determiners for masculine and feminine (Imaz Agirre & Garcia Mayo, 2018, p. 207) in English. The primary focus of this study is the possessive pronouns *his* and *her*, which indicate gender agreement with the possessor in the sentence (see examples (3a) and (3b)).

- (3) a. He’s talking to **his** mother.
 b. She’s talking to **her** mother.

(Imaz Agirre & Garcia Mayo, 2018, p. 207-208)

In English, the possessive pronoun and the definite article are not permitted to co-occur (e.g., **the his mother*). Each of these prepositions functions as a determiner in the sentence, and only one can be employed at a time. Furthermore, they are DPs because pronominals like *her* possess a D feature (Koenenman & Zeijlstra, 2017, p. 102). Consequently, the three languages under investigation exhibit an important difference in their use of the definite article. Hungarian and Romanian necessitate its use with possessive structures, whereas English prohibits it.

2.2. DOM

The term ‘*differential object marking*’ (DOM) was first introduced by Bossong (1985). This term is a comprehensive reference for the cross-linguistic strategies that grammars employ to process specific nominal items, such as nouns or pronouns, in (in)direct object position. This grammatical feature is not present in Hungarian or English, but it is compulsory in Romanian. Clitic pronouns (e.g., *lui* (for)) and dedicated elements (e.g., *pe* (on/onto)) can be employed to express DOM. The current investigation concentrates on the particle *pe*, which serves as both a locative preposition and a DOM-marker in Romanian (Onea, 2023).

There is a degree of agreement among scholars regarding the constraints of DOM usage in Romanian: it is obligatory to use DOM with personal pronouns and proper names, it is possible with modified definite animate NPs and indefinite animate NPs, and it is impossible to use with definite specific and indefinite inanimate NPs (Farkas & von Heusinger, 2003; von Heusinger & Onea, 2008; von Heusinger & Chiriacescu, 2009). The current investigation is focused on obligatory DOM with proper names, as illustrated in example (4):

- (4) Îl caut pe Ion. Romanian
 CL.3SG look for.1SG DOM John
 ‘I am looking for John.’

Once again, in Hungarian and English, this particular grammatical characteristic is not present; nonetheless, it is required in Romanian in the target test sentences (here I refer to the Romanian AJT post-test, where I tested the learners’ sensitivity to grammatical and ungrammatical sentences with regards to DOM, see Appendix 2). (What do you mean by this? Could you please elaborate on/explain what you mean by target test sentences? It would be essential to explain the role of these sentences, how they have been constructed, and why. You have mentioned them in the Appendix, but the process can be understood only if the procedure is explained here in a few lines.

2.3. RCs

A resultative phrase is defined as "an XP that describes the state attained by the referent of the noun phrase it is predicated of as a result of the action named by the verb" (Levin, 1993, p. 101). Romanian does not employ equivalent RC strategies, whereas Hungarian and English do, with some differences discussed later on in this section².

The typology of event lexicalization strategies developed by Talmy (1985, 1991, 2000) distinguishes between satellite-framed languages (e.g., English, German, Dutch, Icelandic, Finnish, Hungarian) and verb-framed languages (e.g., Romance, Japanese, Hebrew). A satellite element that is external to the verb, such as a verbal particle or a resultative secondary predicate, is used to represent result states in English and Hungarian (Kardos & Szávó, 2024). This study focuses on AP RCs, illustrated in example (5):

²Baciu (2014) and Farkas (2009) argue that with some special exceptional cases, Romanian does not exhibit strong resultatives. According to Baciu (2014) it is possible that Romanian has weak resultatives (e.g. pseudo-resultatives). Farkas (2009) argues that ‘Romanian RCs are based only on transitive and intransitive unaccusative matrix verbs’ in contrast with ‘Germanic languages which abound in RCs based on different types of verbs’ (Farkas, 2009, p. 70).

(5) Jones hammered the metal flat.

(Rapoport 1999, p. 669, (38a))

In example (5), the resultative secondary predicate *flat* conveys a result state that is morphologically disassociated from the verb. English and Hungarian are characterized as strong satellite-framed by Acedo-Matellán (2016), in contrast to weak satellite-framed languages, such as Latin (e.g., Romanian) and Slavic languages, which necessitate the expression of results in syntactically autonomous elements that are also prefixed to the verb. However, the word order is a critical distinction between the two languages. In English, resultative phrases are adjuncts and arguably occupy a complement position in the VP, while in Hungarian, they occupy the specifier of an aspectual projection in a preverbal position (Kardos & Farkas, 2022). A Hungarian RC pattern is illustrated in example (6):

(6) Kati laposra kalapált egy vaslemezt. Hungarian
Kati flat.SUBL hammer.PST.3SG an iron_plate.ACC
'Kati hammered an iron plate flat.'

(Kardos & Szávó 2024, p. 108, (18a))

3. The experiment

3.1. Research questions and hypotheses

RQ1: How does transfer from the L1 and L2 occur in the comprehension of pronominal possessors, DOM, and RCs in the initial stages of learning L3 English?

RQ2: What factors determine the choice of transfer source selection? Does transfer happen property-by-property (as argued by the LPM) or wholesale (as predicted by the TPM)?

H1: It is predicted that in the case of all three grammatical conditions, transfer would have non-facilitative effects on the target English grammar (in the initial stages of learning) due to differences between the learned languages with regard to the target structures. Learners may incorrectly believe that a property is shared between the L3 and one or both of the previously acquired languages, which leads to non-facilitative influence (Westergaard et al., 2016, p. 6).

H2: It is anticipated that perceived structural similarity among the languages would drive transfer source selection, predicted by the LPM and TPM. Property-by-property transfer from various sources (L1 and L2) is predicted in line with the LPM (as shown in previous research on pronominal possessors by Imaz Agirre and Garcia Mayo, 2018; on DOM by Giancaspro et al., 2015, and on RCs by Slabakova, 2002, and Whong-Barr, 2005).

Table 1. Predictions of language transfer for the L3 group at the initial stages of L3 English derived from five scenarios.

	Conditions & Transfer source			Type of transfer
	Pronominal Possessors	DOM	RCs	
1. L1 Scenario	Hungarian	-	Hungarian	Facilitative or non-facilitative
2. L2 Status Factor (L2SF)	Romanian	Romanian	-	Facilitative or non-facilitative
3. Cumulative Enhancement Model (CEM)	Romanian	-	-	Facilitative
4. Linguistic Proximity Model (LPM) & Scalpel Model (SM)	Romanian/& Hungarian	Romanian	Hungarian	Facilitative and /or non-facilitative
5. Typological Primacy Model (TPM)	Romanian	Romanian	-	Facilitative or non-facilitative

Table 1 summarises the predictions of five L3 acquisition models of transfer for the L3 group across the three experimental conditions. The L1 Scenario predicts positive or negative transfer solely from the L1, so learners are expected to transfer pronominal possessors and RCs from their L1. However, in the case of DOM, they cannot transfer from their L1 because Hungarian does not have this structure. Evidence of transfer from the L2 would contradict the L1 Scenario. The L2SF predicts facilitative or non-facilitative transfer from the L2 Romanian. However, this cannot be the case with regard to RCs, because Romanian does not allow RCs in the case of the test sentences. The L2SF would be disproved by evidence of transfer from both previously learned languages.

The CEM only hypothesises positive transfer from any or both known languages. Therefore, the model would predict facilitative transfer from the L2 in the pronominal possessor condition, and no transfer for DOM and RCs (as it could only be non-facilitative, which the model does not predict) due to the fact that only Romanian bears some similarities with the English pronominal possessors (e.g. gendered pronominal forms). Evidence of non-facilitation would be against the model. The LPM and SM anticipate property-by-property positive and negative transfer from both L1 and L2 based on actual or perceived structural similarities with L3. Thus, pronominal possessors may transfer non-facilitatively from the L1 and/ or L2, DOM may transfer non-facilitatively from the L2 (the only possible source), and RCs may transfer non-facilitatively from the L1. Evidence of wholesale transfer

(positive or negative) across all three conditions from one of the languages would contradict the models and provide support to the TPM, which predicts the complete initial transfer of the L1 or L2 as a primary hypothesis in the minds of learners upon starting to acquire an L3.

The TPM predicts complete transfer (positive or negative) from the typologically closest language (judged correctly or not). Regarding the target conditions, pronominal possessors and DOM would transfer from Romanian due to some specific linguistic similarity (e.g., gendered pronominal possessors in Romanian and English) and perceived similarity (e.g., obligatory DOM in Romanian, illustrated in ungrammatical test sentences in the English AJT). The model does not predict transfer of RCs from Hungarian, because learners are expected to transfer the Romanian grammar wholesale, as it is structurally more similar to English. Property-by-property transfer for various conditions from both L1 and L2 would be evidence against the TPM.

3.2. Participants and procedure

There were two groups involved in the experiment:

- i) An L3 group – consisting of 24 L1 Hungarian, L2 Romanian, L3 English learners from the fifth grade (11-12 years old) from a school in Romania.
- ii) An L2 group of 20 L1 Hungarian L2 English learners from the fifth grade from a school in Hungary.

Additionally, 12 native adult English speakers from the UK, 1 native Romanian speaker from Romania, 4 native Hungarian speakers from Hungary and Romania helped with validating target sentences and possible translations. The native English speakers also completed the experimental task, performing at a ceiling level.

Learners were chosen based on the following criteria: for the L3 group B1/B2 proficiency level in Romanian, an initial A2 proficiency level in English, and no prior knowledge of any languages other than these three (Hungarian, Romanian, English); for the L2 group A2 proficiency level in English and no prior knowledge of any other languages. All learner participants completed background questionnaires and online placement tests for Romanian (only the L3 group; <https://www.17-minute-languages.com/en/learn-romanian/placement-test/>) and English (Test your English - Young Learners ([cambridgeenglish.org](https://www.cambridgeenglish.org/))) prior to the experiment. Additionally, the L3 group completed a Romanian post-test to ensure that they were familiar with the target structures (pronominal possessors and DOM) in their L2. The specific number of participants from the L3 group whose results were considered varied across conditions (due to their achievement in the Romanian AJT), and their precise number is indicated in the results' interpretation.

After selection, the learners received the link to the Google Forms in which the acceptability judgement task (AJT) was designed. The English experimental AJT and the Romanian post-test (also an AJT in Google Forms) were both completed online in the school's computer laboratory.

3.3. Test instruments

When designing the target test sentences, I considered the potential translations in the learners' L1 and L2 to make sure that the grammatical phenomena under investigation were present in the source languages (Hungarian and/or Romanian). In the case of ungrammatical test sentences, I aimed to test the learners' sensitivity to unacceptable structures. As I mentioned in Section 3.2, I used an English AJT as the experimental test, and a Romanian AJT as a post-test to verify that learners knew the target grammatical phenomena (pronominal possessors and DOM) in their L2. See the short description of the two tests as follows:

i) English AJT with Acceptable/Unacceptable answers, using four experimental conditions (pronominal possessors with Hungarian word order, pronominal possessors with Romanian word order, illustrative DOM, RCs with Hungarian word order) comprised of grammatical and ungrammatical sentence pairs. The total number of test sentences was 50, which consisted of 32 target sentences and 18 unrelated filler sentences (grammatical and ungrammatical pairings). The sentences were presented in the Google Form in a randomized order to prevent learners from copying from each other. Table 2 provides examples of the grammatical and ungrammatical test sentences for each test condition. For a complete list of the test sentences, see Appendix 1.

Table 2. Examples of test sentences (English AJT)

	Conditions	Examples
1.	Pronominal Possessor (Hungarian word order)	John drives Sarah's car because his car is broken. *John drives Sarah's car because the his car is broken.
2.	Pronominal Possessor (Romanian word order)	John helps his mom when she sets the table. *John helps the mom his when she sets the table.
3.	DOM	I understand Maria when she speaks German. * I understand to Maria when she speaks German.
4.	RCs	He burns the toast black . *He black burns the toast.

ii) Romanian AJT for pronominal possessors and DOM conditions. There were 16 target sentences used. Sentences were randomly presented. Table 3 shows examples from the Romanian AJT.

Table 3. Examples of test sentences (Romanian AJT)

	Conditions	Examples
1.	Pronominal Possessor	Sarah folosește stiloul lui Ben pentru că stiloul ei este acasă. Sarah uses pen.DEF he.GEN Ben because pen.DEF she.GEN is.3SG Romanian home 'Sarah uses Ben's pen because her pen is at home.'
2.	DOM	Ion o ajută pe mama lui când aceasta pune masa. John CL. helps DOM mother.DEF he.GEN when this sets table.DEF Romanian 'John helps his mother when she sets the table.'

4. Results

In order to interpret the results, chi-square statistical tests were performed to compare correct and incorrect answers within and between the groups. Accuracy rates are presented in percentages, and figures illustrate the performance of the groups. Results are interpreted in relation to L3 acquisition models, with special focus on transfer effects from previously learned languages. The following subsections discuss the results of each test condition separately.

4.1. Pronominal possessors

In order to accommodate potential transfer from both L1 Hungarian and L2 Romanian, the experimental condition of pronominal possessors was divided into two sections. The word order of Hungarian and Romanian was replicated by constructing ungrammatical sentences. It is essential to note that the total number of L3 participants for this condition was 16 out of 24, as 8 learners were eliminated based on the selection criteria (they were unable to demonstrate knowledge of pronominal possessors in L2 Romanian in the post-test). Pronominal possessors were expected to transfer non-facilitatively from the L1 or the L2 due to structural and word order differences.

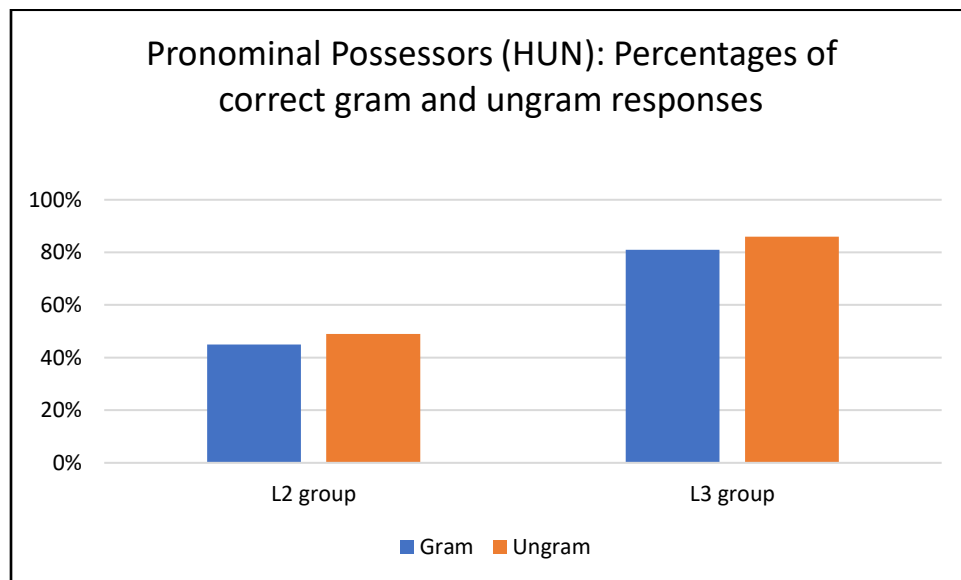
4.1.1. Hungarian word order

The L2 group demonstrated low overall accuracy, with 47% of their answers correct and 53% incorrect. They performed slightly better at rating ungrammatical sentences (49% correct answers) than grammatical sentences (45% correct answers). However, the chi-square statistical test did not produce a significant result (chi-square statistic: 0.2259, the p-value is .634594, not significant at $p < .05$). The results may indicate nonfacilitative transfer effects from the learners' L1 Hungarian due to the lack of natural and grammatical gender and the obligatory use of the definite article in Hungarian. If we examine the ungrammatical test sentences in detail, we find that

over half of the learners accepted the ungrammatical sentences with pronominal possessors preceded by the definite article (copying the Hungarian structure) in three out of four sentences. The results of the L2 group also show that they have not yet fully set the parameters of their L2 English in the case of pronominal possessors at this level of proficiency (A2).

The L3 group (N = 16) exhibits high accuracy rates, with 84% of their answers being correct and only 16% incorrect. Similarly to the L2 group, they performed slightly better at rating ungrammatical sentences (86% correct answers) than grammatical sentences (81%). However, the difference is not statistically significant (chi-square statistic: 0.5127, p-value = 0.473979, not significant at $p < 0.05$). The highly accurate results may indicate facilitative transfer from the learners' L2, Romanian, due to some similarities between Romanian and English, such as the presence of gendered pronominal possessors. Due to their probable frequent exposure to the grammatical structure, the learners may have identified this similarity between their L2 and L3 at an early stage and learned to inhibit non-facilitative transfer effects of their L1. The chi-square test yielded a statistically significant result (chi-square statistic: 41.2212, p-value < 0.00001 , significant at $p < 0.05$) when comparing the correct and incorrect answers between the two groups, indicating a considerable difference between them, with the L3 group being more accurate. Figure 1 shows the proportions of correct answers of the two groups in the case of grammatical and ungrammatical test sentences.

Figure 1. Percentages of correct responses for Pronominal Possessors – Hungarian patterned condition



4.1.2. Romanian word order

The L2 group's accuracy rate was 53% correct and 47% incorrect answers. They performed slightly better at rating grammatical sentences (66% correct answers) than ungrammatical sentences (40% correct answers). The chi-square statistical test gave significant result when comparing the number of correct and incorrect grammatical and ungrammatical answers (chi-square statistic: 11.0682, the p-value is .000878, significant at $p < .05$). The results may indicate non-facilitative transfer effects from the learners' L1 Hungarian due to the lack of natural and grammatical gender and the obligatory use of the definite article in Hungarian. Their error rates are high, similar to the results with the Hungarian-patterned test sentences. The overacceptance of ungrammatical sentences is likely a result of low proficiency in L2 English.

The accuracy rate of the L3 group was 74% correct and 26% incorrect answers, much higher than the L2 group's rates. Similarly to the results of the L2 group, L3s performed slightly better at rating grammatical sentences (78%) than ungrammatical sentences (70%). However, upon comparison of correct and incorrect answers, the chi-square statistical test did not produce a significant result (chi-square statistic: 1.0207, p-value = 0.312345, not significant at $p < 0.05$).

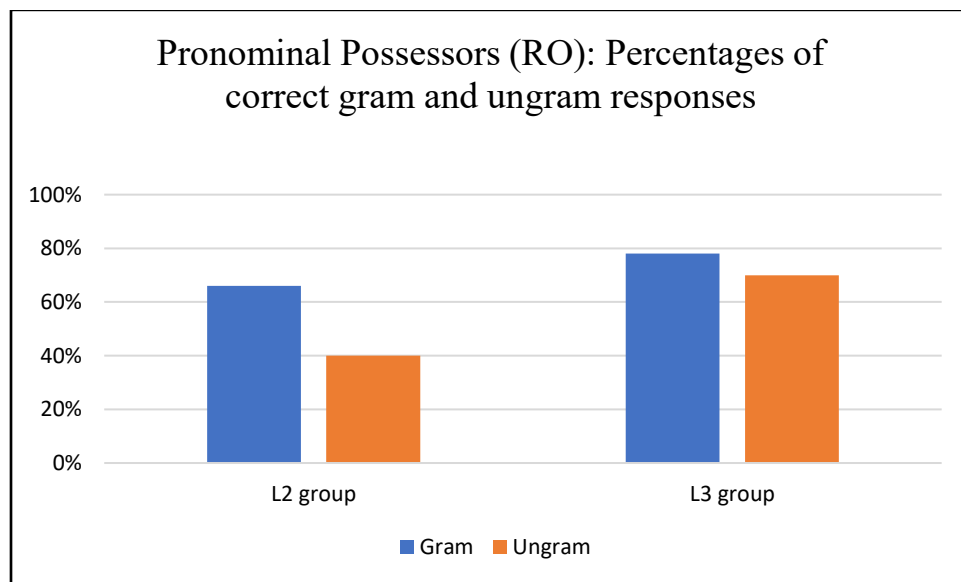
Contrary to the predictions, learners did not overaccept ungrammatical sentences with the Romanian pronominal possessor pattern, so it is highly probable that there was no significant non-facilitation of the L2 in this case. It is reasonable to assume that the L2 had a facilitative effect due to structural similarity, similar to the ratings of the ungrammatical sentences with a Hungarian pattern. The high accuracy rates indicate that L3 learners have already set the parameters of the target L3 English grammar, with regard to the ungrammaticality of using the definite article with pronominal possessors and the correct word order, which differs from their L1 and L2. Also, the learners may have identified and exploited the similarity between the gendered pronominal possessor forms that agree with the possessee in Romanian and English during the experimental testing. Concerning the acquisition of pronominal possessors in the literature, in a study Sólyom (2023) looked at animacy and gender agreement effects in the case of the same language combination (L1 Hungarian L2 Romanian and L3 English) and the results indicated that natural gender of the head nouns helped learners establish correct gender agreement with the possessee. It can be assumed that in the present study the learners could identify this structural similarity regarding the role of natural gender and agreement in pronominal possessor forms, the word order differences between the L2 and L3 and they could also inhibit the non-facilitative effect of the L1 and L2 related to the use of the definite article.

The chi-square test yielded a statistically significant result (chi-square statistic: 13.5, $p < 0.05$) when comparing the correct and incorrect answers between the two

groups, indicating a considerable difference between them, with the L3 group being more accurate. The data show evidence for facilitative effects of the L2 Romanian. Thus, the overall results of the pronominal possessors (Hungarian and Romanian patterned) condition are in line with the predictions of the SM, and LPM in the sense that learners are capable of transferring properties from the structurally more similar language and separating them with scalpel-like precision (as predicted by Slabakova, 2017).

Figure 2 illustrates the accuracy rates of the two groups with regard to grammatical and ungrammatical sentences with pronominal possessors with Romanian word order in the ungrammatical test sentences.

Figure 2. Percentages of correct responses for Pronominal Possessors – Romanian patterned condition



4.2. DOM

The L2 group's accuracy rate was 49% correct and 51% incorrect answers. They performed slightly better at rating grammatical sentences (59% correct answers) than ungrammatical sentences (39% correct answers). The chi-square statistical test yielded a significant result (chi-square statistic = 6.404; p -value = 0.011386; significant at $p < 0.05$). The L2 group's low accuracy can be attributed to their limited proficiency in English and the lack of structural similarity with regard to DOM structures in Hungarian.

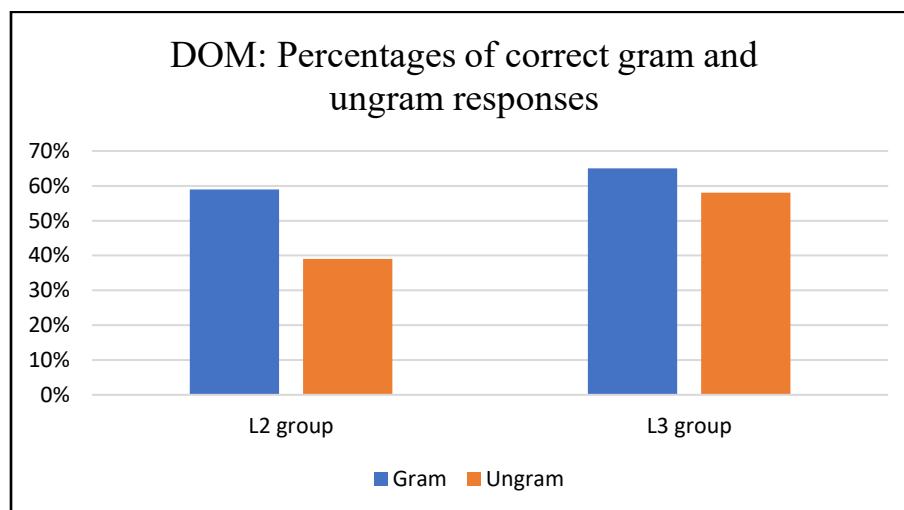
The L3 group's ($N = 15$) results also show low accuracy levels: 61% correct and 39% incorrect answers. Similarly to the L2 group, they performed slightly better at rating grammatical sentences (65% correct answers) than ungrammatical sentences (58% correct answers). However, the chi-square statistical test did not produce a

significant result (chi-square statistic: 0.564, the p-value is 0.452636, not significant at $p < 0.05$). The results of the L3 group may possibly indicate non-facilitative effects of the L2 Romanian, where DOM is obligatory in the case of the test sentences. Due to high exposure and frequency of the structure in Romanian, learners at the initial stages of L3 learning might have perceived wrong structural similarity and transferred DOM property into their target L3 English (acceptance of 42% of the ungrammatical test sentences with artificial DOM in English may lend support to this hypothesis).

The results of the English AJT, paired with those of the Romanian AJT regarding the availability of DOM for transfer in the learners' L2, provide significant evidence for potential non-facilitative transfer from the L2 Romanian, in line with the predictions of the SM, LPM, TPM, and L2SF. To cite exemplary research that yielded comparable outcomes, Giancaspro et al. (2015) conducted a study on the acquisition of DOM in L3 Brazilian Portuguese (BP) by L2s who simultaneously spoke Spanish and English. The study examined Spanish/English successive L2s and Spanish/English heritage speakers. They employed a grammaticality judgement task (GJT) in both written and oral formats. In both groups, they observed non-facilitative transfer from Spanish, which they interpreted as supporting evidence for the TPM (Rothman 2010, 2011, 2013), indicating that learners transferred the most structurally/typologically similar language during the initial stages of L3 acquisition.

The chi-square test gave statistically significant result (chi-square statistic: 4.2708, the p value is .038772, significant at $p < .05$) when comparing the two group's correct and incorrect answers, showing that there is a considerable difference between the answers of the two groups, the L3 group being more accurate here, too. Figure 3 illustrates the accuracy rates of the two groups with regard to DOM condition.

Figure 3. Percentages of correct responses for DOM condition



The results of the L2 and L3 groups might seem contradictory because one would expect the L2s to reject the ungrammatical test sentences to a greater degree. After all, neither their L1 Hungarian nor their L2 English allows DOM structures. Learners of L2 English may know that a lack of DOM is grammatical without knowing that the presence of (illustrative) DOM is ungrammatical in English. Thus, instances of accepting ungrammatical test sentences may be evidence of linguistic development in the case of the L2 group (see a similar finding regarding DOM in Cabrelli et al. 2020). Moreover, the L3 group's superior performance in all test conditions may be attributed to the extensive evidence of the metalinguistic advantage of multilingual learners, whose knowledge of more grammatical systems makes learning a new language easier (e.g., Bialystok, 2001).

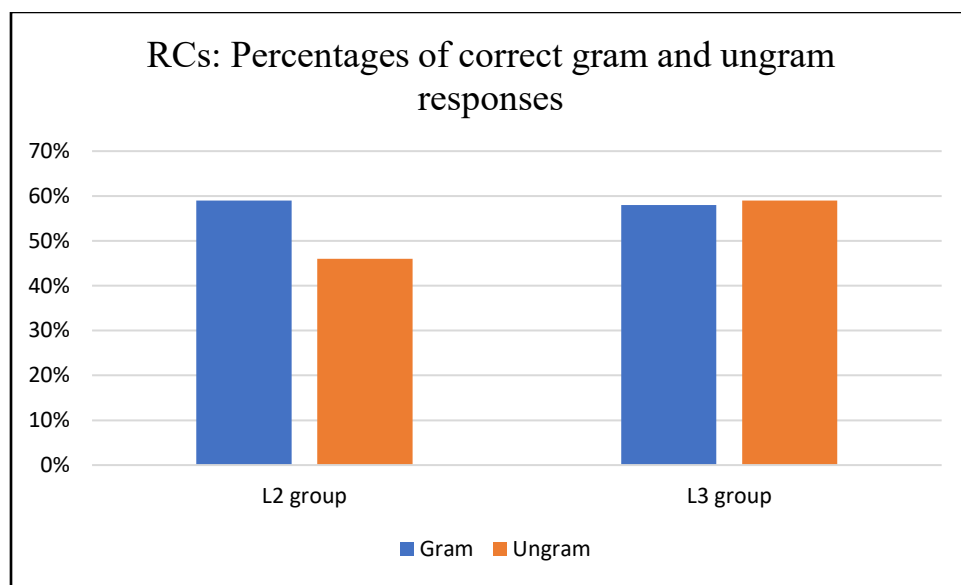
4.3. RCs

The L2 group's accuracy rate was 53% correct and 47% incorrect answers. Learners performed better at rating grammatical sentences (59% correct answers) than ungrammatical sentences (46%). However, the difference is not statistically significant (the chi-square statistic is 2.5063, and the p -value is 0.113394, which is not significant at $p < 0.05$). They accepted the ungrammatical sentences that copied Hungarian RC word order in 54% of the answers. This low accuracy could be explained by non-facilitative transfer effects from the L1 Hungarian, where RCs appear in different word order. Also, they have limited exposure to the structure in their L2 English at low proficiency levels.

The L3 group's ($N = 24$) accuracy is similar to the L2s': 59% correct and 41% incorrect answers. They performed very similarly when rating grammatical (58% accurate) and ungrammatical sentences (59% accurate), and the chi-square statistic did not give statistically significant result (the chi-square statistic is 0.0215, the p -value is .883404, not significant at $p < .05$). The learners accepted the ungrammatical sentences that copied Hungarian RC word order in 41% of the answers. Therefore, it is possible that L2 learners also resorted to their L1, as they might have perceived structural similarities in the case of RCs in Hungarian and English. However, due to word order differences, this perception resulted in non-facilitative transfer from the L1 Hungarian. Similarly to the L2 group, they have limited exposure to the structure in English at low proficiency levels (and RCs are never explicitly taught in schools), and they could not turn to their L2 because there are no RCs in Romanian (with some exceptional cases not used in the test sentences of the experiment). The chi-square test did not yield a statistically significant result (chi-square statistic = 1.4298, p -value = 0.113394, not significant at $p > 0.05$) when comparing the two groups' correct and incorrect answers, indicating that there is no significant difference between the two groups.

The results of the RCs condition are compatible with the predictions of the LPM, and also confirm the study's hypothesis that predicted non-facilitative property-by-property transfer from the learners' L1 due to perceived (and actual) structural similarity between Hungarian and English. The entirety of the results (from all three conditions) appears to be more compatible with the predictions of the LPM, as learners transfer from various sources (both L1 and L2) both facilitatively and non-facilitatively, suggesting property-by-property transfer. Figure 4 shows the accuracy rates of the two groups regarding the RCs condition.

Figure 4. Percentages of correct responses for RCs condition



5. Conclusions

This paper aimed to examine the role of transfer in the development of two groups: L2 learners (L1 Hungarian, L2 English) and L3 learners (L1 Hungarian, L2 Romanian, L3 English), at the initial stages of learning English. The two groups were age- and proficiency-matched from two different countries: Hungary and Romania. By comparing the two groups, it was possible to locate both the source and the nature of linguistic transfer, especially in the case of the L3 group (which was the primary focus of the study). The results show that the L3 group significantly outperformed the L2 group in two out of the three test conditions. This result may show evidence of the metalinguistic advantage of the L3 group, stemming from their knowledge of more than one previously learned language, as well as increased cognitive performance and awareness.

Regarding the first research question and hypothesis, the predictions were partly confirmed. It was predicted that the L3 group would only show non-facilitative

transfer effects due to differences between the grammatical structures. However, the L3 group's accuracy (over 80%) pointed to a possible significant facilitative impact due to some similarities between the L2 Romanian and L3 English in the case of pronominal possessor condition (which similarities were not considered in the initial predictions). The learners may have realised the presence of gendered pronominal possessor forms in both languages, which helped them to provide correct answers in the English AJT. Turning to the second research question, in the case of the pronominal possessor condition, the determining factor for transfer source selection was actual structural similarity between the L2 and L3. Four of the five transfer models specified in Table 1 (see Section 3.1) are compatible with facilitative transfer from the L2, with the exception of the L1 Scenario.

Confirming the hypothesis, the L3 group demonstrated non-facilitative transfer effects from the L2 Romanian in DOM condition. They were unable to reject the ungrammatical test sentences with illustrative DOM in the English AJT. Perceived structural similarity between the L2 and L3 might have driven their choice; additionally, they could not use their L1 because Hungarian does not have DOM structures. Results are compatible with the L2SF, LPM, SM, TPM, and contra the L1 Scenario (with evidence of L2 transfer) and the CEM (with evidence of negative transfer).

In RCs condition, the hypothesis was confirmed as the data from the L3 group presents significant non-facilitative transfer effects from L1 Hungarian. Actual structural similarity between Hungarian and English RCs might have driven the selection of the transfer source language; however, word order differences resulted in non-facilitation (as learners were unable to reject ungrammatical sentences with Hungarian RC word order). The data provides evidence for the L1 Scenario (but only for the RC condition), LPM, SM (non-facilitative transfer from the L1) and contra the L2SF (with transfer from the L1), CEM (with negative transfer), and TPM (which would not predict transfer from the L1 after wholesale transfer of the L2 – as discussed in Section 3.1).

The entirety of the data regarding all three conditions provides significant evidence of both positive and negative CLI and it can be concluded that the LPM and the SM can account for most of the results, with property-by-property facilitative transfer (for pronominal possessor condition) and non-facilitative transfer (for DOM and RCs conditions) from the L1 and the L2. Perceived linguistic proximity was the driving force behind transfer source selection (except for the pronominal possessor condition). An unexpected finding of the study was evidence for metalinguistic advantage of multilinguals (Bialystok, 2001), as the L3 group had higher accuracy rates in the English comprehension task (AJT) in all test conditions than the L2 group (for pronominal possessors and DOM the difference was statistically significant

between the two groups)³. The combination of languages (Hungarian, Romanian, English) from different language families (Finno-Ugric, Romance, Germanic) and the analyses of three target grammatical features within a single experiment make a substantial contribution to the field of linguistics, particularly in the field of L3 acquisition and language teaching in general.⁴ This analysis provides insight into learners' processing, with a special focus on linguistic transfer during the early stages of learning an L3.

References

- Acedo-Matellán, V. (2016). *The Morphosyntax of Transitions: A Case Study in Latin and Other Languages*, Oxford Studies in Theoretical Linguistics (Oxford, 2016; online edn, Oxford Academic, <https://doi.org/10.1093/acprof:oso/9780198733287.001.0001>)
- Baciu, I. (2014). Towards a Typology of Resultative Constructions: A Romanian Perspective. In R. Cosma, S. Engelberg, S. Schlotthauer, S. Stanescu & G. Zifonun (Ed.), *Komplexe Argumentstrukturen: Kontrastive Untersuchungen zum Deutschen, Rumänischen und Englischen* (pp. 197-226). Berlin, München, Boston: De Gruyter. <https://doi.org/10.1515/9783110343229.197>
- Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, 23(4), 459–484. <https://doi.org/10.1177/0267658307080557>
- Bialystok, E. (2001). *L2ism in Development: Language, Literacy, and Cognition*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511605963>
- Bossong, G. (1985). *Empirische Universalienforschung*. Differentielle Objektmarkierung in den neuiranischen Sprachen. Tübingen, Narr
- Cabrelli, J., Iverson, M., Giancaspro, D., & González, B. H. (2020). The roles of L1 Spanish versus L2 Spanish in L3 Portuguese morphosyntactic development. *Issues in Hispanic and Lusophone linguistics* 11–33. <https://doi.org/10.1075/ihll.24.01cab>
- De Angelis, G. (2007). *Third or Additional Language Acquisition*. Bristol, Blue Ridge Summit: Multilingual Matters. <https://doi.org/10.21832/9781847690050>
- Dobrovie-Sorin, C. & Giurgea, I. (eds.) (2013). *A Reference Grammar of Romanian: Volume 1: The noun phrase*. Amsterdam, Philadelphia: John Benjamins. <https://doi.org/10.1075/la.207>
- É. Kiss, K. (2002). *The Syntax of Hungarian* (Cambridge Syntax Guides). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511755088>
- Falk, Y., & Bardel, C. (2011). Object pronouns in German L3 syntax: Evidence for the L2 Status Factor. *Second Language Research*, 27(1), 59–82. <https://doi.org/10.1177/0267658310386647>
- Farkas, D. & von Heusinger, K. (2003). Stability of reference and object marking in Romanian. Paper presented at Workshop on Direct Reference and Specificity, ESSLLI, Vienna, August 2003.
- Farkas, I. Á. (2009). Some differences between English and Romanian resultative structures. *Bucharest Working Papers in Linguistics XI/2*, 59-71.

³ Bialystok (2001) contends that bilingual children may develop executive functions and the control over attention and inhibition in a distinctively and more advantageous manner due to the constant management of two competing languages.

⁴ A reviewer commented that both Romance and Germanic language families are sub-branches of the Indo-European language family, and there is a strong grammatical influence on English coming from Latin and French, and Romanian also has a Latin grammatical core structure. If I understand it well, the reviewer suggests that the languages I work with do not necessarily belong to different language families. However, I contend that even though there are common influences from Latin, the grammatical phenomena under investigation differ in the three languages, and therefore they can be used for examining transfer and testing L3 models of language acquisition.

- Flynn, S., Foley, C., & Vinnitskaya, I. (2004). The Cumulative-Enhancement model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition. *International Journal of Multilingualism*, 1(1), 3–16. <https://doi.org/10.1080/14790710408668175>
- Giancaspro, D., Halloran, B., & Iverson, M. (2015). Transfer at the initial stages of L3 Brazilian Portuguese: A look at three groups of English/Spanish L2s. *L2ism: Language and Cognition*, 18, 191–207. <https://doi.org/10.1017/S1366728914000339>
- Hermas, A. (2010). Language acquisition as computational resetting: verb movement in L3 initial state." *International Journal of Multilingualism* 7: 202–24. <https://doi.org/10.1080/14790718.2010.487941>
- Hermas, A. (2015). The categorization of the relative complementizer phrase in third-language English: A feature re-assembly account. *International Journal of L2ism* 19: 587–607. <https://doi.org/10.1177/1367006914527019>
- Imaz Agirre, A. & García Mayo, M.D.P. (2018). Proficiency and transfer effects in the acquisition of gender agreement by L2 and L3 English learners. *Meaning and Structure in Second Language Acquisition: In honor of Roumyana Slabakova*, Edited by Jacee Cho, Michael Iverson, Tiffany Judy, Tania Leal and Elena Shimanskaya, *Studies in L2ism* (55). 203-227. <https://doi.org/10.1075/sibil.55>
- Kardos, É., & Farkas, I.-Á. (2022). The syntax of inner aspect in Hungarian. *Journal of Linguistics*, 58(4), 807–845. <https://doi.org/10.1017/S0022226721000426>
- Kardos, É. & Szávó, A. (2024). Event lexicalization in Hungarian. In M. den Dikken & H. Kishimoto (eds.), *Formal Perspectives on Secondary Predication*, 95-126. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110981742-004>
- Koenenman, O., & Zeijlstra, H. (2017). *Introducing Syntax*. Cambridge University Press. <https://doi.org/10.1017/9781316156391>
- Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press. <https://doi.org/10.2307/415968>
- Odlin, T. (2003). *Language transfer: Cross-linguistic influence in language learning*, Cambridge University Press. <https://doi.org/10.1017/CBO9781139524537>
- Onea, E. (2023). Virginia Hill and Alexandru Mardale: The diachrony of Differential Object Marking in Romanian. *Folia Linguistica*, vol. 57, no. 1, pp. 227-241. <https://doi.org/10.1515/flin-2023-2003>
- Ranong, S. & Leung, Y. (2009). Chapter 8. Null Objects in L1 Thai–L2 English–L3 Chinese: An Empiricist Take on a Theoretical Problem. In Y. Leung (Ed.), *Third Language Acquisition and Universal Grammar* (pp. 162-191). Bristol, Blue Ridge Summit: Multilingual Matters. <https://doi.org/10.21832/9781847691323-011>
- Rothman, J., & Cabrelli Amaro, J. (2010). What variables condition syntactic transfer? A look at the L3 initial State. *Second Language Research*, 26, 189–218. <https://doi.org/10.1177/0267658309349410>
- Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The Typological Primacy Model. *Second Language Research*, 27, 107–127. <https://doi.org/10.1177/0267658310386439>
- Rothman, J. (2013). Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance. In S. Baauw, F. Dirjkoningen, & M. Pinto (Eds.), *Romance languages and linguistic theory 2011* (pp. 217–247). Amsterdam: John Benjamins. <https://doi.org/10.1075/rllt.5.11rot>
- Rothman, J. (2015). Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *L2ism: Language and Cognition*, 18(2), 179–190. <https://doi.org/10.1017/S136672891300059X>
- Slabakova, R. (2002). The Compounding Parameter in Second Language Acquisition. *Studies in Second Language Acquisition* 24 (4), <https://doi.org/10.1017/S0272263102004011>

- Slabakova, R. (2017). The Scalpel Model of third language acquisition. *International Journal of L2ism*, 21(6), 651–666. <https://doi.org/10.1177/1367006916655413>
- Sólyom, Cs. I. (2023). L3 Acquisition of Pronominal Possessors and Gender Agreement. *Argumentum*, 19, 156–182. Debreceni Egyetemi Kiadó. <https://doi.org/10.34103/ARGUMENTUM/2023/8>
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (ed.), *Language Typology and Syntactic Description 3: Grammatical Categories and the Lexicon*, 57–149. Cambridge: Cambridge University Press.
- Talmy, L. (1991). Path to realization: A typology of event conflation. In Laurel A. Sutton, Christopher Johnson & Ruth Shields (eds.), *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society*, 480–519. Berkeley, CA: Berkeley Linguistics Society.
- Talmy, L. (2000). *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/6847.001.0001>
- von Heusinger, K. & Onea E.G. (2008). Triggering and blocking effects in the diachronic development of DOM in Romanian. *Probus* 20, 67–110. <https://doi.org/10.1515/PROBUS.2008.003>
- von Heusinger, K. & Chiriacescu, S. (2009). Definite “bare” nouns and pe-marking in Romanian. In Espinal, M.T., Leonetti, M. & McNally, L. (eds.), *Proceedings of the IV Nereus International Workshop on “Definiteness and DP Structure in Romance Languages.” Arbeitspapier* 124, 63–82. Fachbereich Sprachwissenschaft, Universität Konstanz.
- Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of L2ism*, 21(6), 666–682. <https://doi.org/10.1177/1367006916648859>
- Westergaard, M. (2020). Microvariation in Multilingual Situations: The Importance of Property-by-Property Acquisition. *Second Language Research*. <https://doi.org/10.1177/0267658319884116>
- Whong-Barr, M. (2005). Morphology, Derivational Syntax and Second Language Acquisition of Resultatives. Unpublished Doctoral dissertation. Durham University, United Kingdom.

Appendix 1

Test sentences for the English AJT

Target sentences

i) Pronominal Possessor - Romanian word order

- 1 Brian is eating from Emily's lunchbox because his lunchbox is empty.
- *Brian is eating from Emily's lunchbox because the lunchbox his is empty⁵.
- 2 Sarah uses Ben's pen because her pen is at home.
- *Sarah uses Ben's pen because the pen her is at home.
- 3 Sue is having lunch and she asks her dad for some bread.
- *Sue is having lunch and she asks the dad her for some bread.
- 4 John helps his mom when she sets the table.
- *John helps the mom his when she sets the table.

ii) Pronominal possessor - Hungarian word order

- 1 Emma borrows Tom's phone because her battery is running low.
- *Emma borrows Tom's phone because the her battery is running low.
- 2 Mike is reading Lisa's book because his copy is missing.
- *Mike is reading Lisa's book because the his copy is missing.
- 3 Rachel is staying at a friend's house because her apartment is being painted.
- *Rachel is staying at a friend's house because the her apartment is being painted.
- 4 John drives Sarah's car because his car is broken.
- *John drives Sarah's car because the his car is broken.

iii) DOM

- 1 Paul visits Sam every Sunday.
- *Paul visits to Sam every Sunday.
- 2 Mike hugs Nelly when they meet at school.
- *Mike hugs to Nelly when they meet at school.
- 3 I understand Maria when she speaks German.
- * I understand to Maria when she speaks German.
- 4 I know Cory from kindergarten.
- *I know to Cory from kindergarten.

iv) RCs

- 1 Grandma wipes the dirty table clean.
- *Grandma clean wipes the dirty table.
- 2 Danny hammers the metal flat.

⁵ The * symbol marks the ungrammatical sentences.

- * Danny flat hammers the metal.
- 3 He burns the toast black.
- *He black burns the toast.
- 4 Meggy sweeps the kitchen floor clean twice a week.
- *Meggy clean sweeps the kitchen floor twice a week.

Filler sentences

- 1 The water is warm at the swimming pool.
- *The water is warm the at swimming pool.
- 2 There are some offices on the fifth floor.
- *There are some office on the fifth floor.
- 3 We usually have dinner at 8.
- *We have usually dinner at 8.
- 4 Ben's bike has got a flat tire.
- *Ben's bike got has a flat tire.
- 5 My grandma can cook really well.
- *My grandma can cooks really well.
- 6 I tidy my room every day.
- *I my room tidy every day.
- 7 Megan likes horse-riding very much.
- *Megan very likes horse-riding.
- 8 There is a lot of water on the carpet.
- *There a lot of water is on the carpet.
- 9 Dan plays the piano very often.
- *Dan play the piano very often.

Example of a test sentence as it appeared in the AJT Google Form:

*Please read each sentence carefully and under each sentence select the option that you consider correct: **Acceptable** (if it feels correct in English) or **Unacceptable** (if it feels incorrect in English).*

Grandma wipes the dirty table clean.

- ☐ Acceptable
- ☐ Unacceptable

Appendix 2

Test sentences for the Romanian AJT

Target sentences

i) Pronominal Possessor - Romanian word order

1 Brian mănâncă din cutia Emiliei, deoarece cutia lui de prânz este goală.
 Brian eats from box.DEF Emily.GEN because box.DEF he.GEN of lunch is empty
 ‘Brian eats from Emily’s box, because his lunchbox is empty.’

*Brian mănâncă din cutia Emiliei, deoarece lui cutia de prânz este goală.
 Brian eats from box.DEF Emily.GEN because he.GEN box.DEF of lunch is empty
 ‘Brian eats from Emily’s box, because his lunchbox is empty.’

2 Sarah folosește stiloul lui Ben pentru că stiloul ei este acasă.
 Sarah uses pen.DEF he.GEN Ben because pen.DEF she.GEN is home
 ‘Sarah uses Ben’s pen because her pen is at home.’

*Sarah folosește stiloul lui Ben pentru că ei stiloul este acasă.
 Sarah uses pen.DEF he.GEN Ben because she.GEN pen.DEF is home
 ‘Sarah uses Ben’s pen because her pen is at home.’

3 Susana ia prânzul și îi cere tatălui ei niște pâine.
 Susana has lunch.DEF and CL.3SG asks father.DEF she.GEN some bread
 ‘Susana is having lunch and asks her father for some bread.’

*Susana ia prânzul și îi cere ei tatălui niște pâine.
 Susana has lunch.DEF and CL.3SG asks she.GEN father.DEF some bread
 ‘Susana is having lunch and asks her father for some bread.’

4 Ion o ajută pe mama lui când aceasta pune masa.
 John CL.3SG.F helps to mother.DEF he.GEN when this.F sets table.DEF
 ‘John helps her mother when she sets the table.’

*Ion o ajută pe lui mama când aceasta pune masa.
 John CL.3SG.F helps to he.GEN mother.DEF when this.F sets table.DEF
 ‘John helps her mother when she sets the table.’

ii) DOM

1 Paul îl vizitează pe Samuel în fiecare duminică.
 Paul CL.3.SG.M visit DOM Samuel in every Sunday
 ‘Paul visits Samuel every Sunday.’

*Paul îl vizitează Samuel în fiecare duminică.
 Paul CL.3.SG.M visit Samuel in every Sunday
 ‘Paul visits Samuel every Sunday.’

2 Nicolae o îmbrățișează pe Nora când se întâlnesc la școală.
 Nicolae CL.3SG.F hugs DOM Nora when CL.3PL meet at school
 ‘Nicolae hugs Nora when they meet at school.’

*Nicolae o îmbrățișează Nora când se întâlnesc la școală.
 Nicolae CL.3SG.F hugs Nora when CL.3PL meet at school
 ‘Nicolae hugs Nora when they meet at school.’

3 O înțeleg pe Maria când vorbește germană.
 CL.3SG understand.1SG DOM Maria when speak.3SG German
 ‘I understand Maria when she speaks German.’

*O înțeleg Maria când vorbește germană.
 CL.3SG.F understand.1SG Maria when speak.3SG German
 ‘I understand Maria when she speaks German.’

4 O cunosc pe Corina de la grădiniță.
 CL.3SG.F know DOM Corina from nursery
 ‘I know Corina from nursery.’

*O cunosc Corina de la grădiniță.
 CL.3SG.F know Corina from nursery
 ‘I know Corina from nursery.’

ANNAMÁRIA KISS¹ – VALÉRIA CSÉPE^{2,3}

¹Multilingualism Doctoral School, University of Pannonia

annamaria.kiss@phd.uni-pannon.hu

<https://orcid.org/0009-0009-1194-0341>

²HUN-REN RCNS Brain Imaging Centre

³University of Pannonia

csepe.valeria@ttk.hu

<https://orcid.org/0000-0002-5021-6024>

Annamária Kiss – Valéria Csépe: Investigating working memory impairments in bilingual aphasia

Alkalmazott Nyelvtudomány, XXV. évfolyam, 2025/1. szám, 161–186.

doi:<http://dx.doi.org/10.18460/ANY.2025.1.009>

Investigating working memory impairments in bilingual aphasia

While most aphasia studies investigate working memory impairments in monolinguals, there is a lesser focus on bilinguals. This study aimed to compare the assumed working memory impairments in individuals with monolingual and bilingual aphasia. Therefore, the study employed both visual and verbal working memory tasks, including the Visual Sequential Memory Test, the Rey-Osterrieth Complex Figure B, and the Auditory Sequential Memory Test. It was hypothesised that both aphasia groups would exhibit impairments in their visual and verbal working memory performance. A total of 54 participants were recruited, including monolingual (N=10) and bilingual (N=8) individuals with aphasia as well as individuals without any neurological problems (N=36). Both aphasia groups demonstrated significant variability in the visual and auditory memory tasks, with impairments noted in comparison to the reference groups. Although the bilingual aphasia group had lower scores than the monolinguals, these differences were not statistically significant. These findings contribute to our understanding of the less explored domain of working memory impairments in bilingual aphasia.

Keywords: aphasia, bilingual aphasia, working memory impairment, visual and auditory working memory, multiple case study

1. Introduction

Aphasia, a language disorder resulting from brain damage, presents unique challenges in understanding the cognitive processes underlying language impairment. Although extensive research has focused on working memory deficits in monolingual individuals with aphasia, less attention has been paid to those shown by bilingual aphasics. Therefore, there is a significant gap in the relevant literature regarding the characteristics of their working memory, a system responsible for temporal holding and manipulating information, which is crucial for language processing and communication. This study aims to bridge this gap by providing data acquired by testing working memory in monolingual and bilingual individuals with aphasia.

1.1. Aphasia

Aphasia is regarded as an acquired language disorder following brain damage, leading to impairments in all language modalities (Chapey, 2008). According to the multidimensional view, multiple forms of aphasia exist, all influenced by and corresponding to distinct underlying lesion sites that exhibit individual characteristics (Goetz, 2007). A broadly accepted distinction is made between fluent (Wernicke's, conduction, transcortical sensory, and anomic) and non-fluent (Broca's, global, transcortical motor, and mixed transcortical) subtypes of aphasia. The main characteristics of subtypes (*non-fluent with Italic*) of aphasia are presented in Table 1.

Table 1. Main characteristics of subtypes of aphasia

Aphasia type	Speech	Comprehension	Repetition	Naming
Wernicke's	fluent, empty	poor	poor	poor
Conduction	fluent	good	poor	poor
Transcortical sensory	fluent	poor	good	poor
Anomic	fluent with circumlocutions	good	good	poor
<i>Broca's</i>	<i>poor</i>	<i>good</i>	<i>poor</i>	<i>poor</i>
<i>Global</i>	<i>extremely poor</i>	<i>poor</i>	<i>poor</i>	<i>poor</i>
<i>Transcortical motor</i>	<i>little</i>	<i>good</i>	<i>good</i>	<i>poor</i>
<i>Mixed transcortical</i>	<i>extremely poor</i>	<i>poor</i>	<i>good</i>	<i>poor</i>

1.2. Bilingual Aphasia

Bilingual aphasia is a condition in which an individual who is fluent in two or more languages experiences language deficits in both languages following brain damage. This condition presents unique challenges for assessment, diagnosis, and treatment due to the complex interaction between the two languages in the brain. Bilingual aphasia can be classified into two types of impairments: parallel and selective. Parallel impairment occurs when both languages of a bilingual person are affected to a similar extent. Selective impairment, on the other hand, refers to a situation where one language is more affected than the other (Gitterman et al., 2012).

1.3. Working memory in aphasia

Working memory (WM) is defined as a system of limited capacity that is responsible for holding and manipulating information while undertaking cognitive tasks (Baddeley, 2003). Most studies investigating the working memory of individuals with monolingual aphasia have consistently shown an impairment in updating (Murray et al., 2001; Wright & Shisler, 2005; Potagas et al., 2011; Mayer & Murray, 2012). However, it is notable that many studies have employed linguistically demanding tasks. To complete these tasks, individuals must engage in semantic and/or phonological processing, enabling them to understand the instructions and formulate an appropriate response. Some studies have focused on investigating non-linguistic working memory, with particular attention paid to ensure the tasks were simple and fully non-linguistic (Christensen & Wright, 2010; Christensen et al., 2018)

Christensen and Wright (2010) conducted a study examining both linguistic and non-linguistic working memory in a sample of 12 monolingual individuals with aphasia whose language impairments varied as measured by the Aphasia Quotient (AQ) of the Western Aphasia Battery (WAB), with scores ranging from 47.9 to 90.7. The researchers employed the *n*-back task using three distinct categories of stimuli, different in terms of linguistic load. This linguistic load refers to the semantic and phonological variations of the stimuli. The category “fruit” represented the highest linguistic load, fribbles, e.g., novel objects could be classified as semi-linguistic while blocks were classified as non-linguistic stimuli. The Christensen and Wright study (2010) suggested that the working memory performance was significantly improved by verbal encoding, especially in aphasic participants. Their findings indicate that aphasic participants demonstrated a similarity in performance with the control group regarding WM tasks that varied according to linguistic load, though the performance of aphasics was limited. It was also observed that decreasing the linguistic elements of the task resulted in a decrease in the performance accuracy of aphasic participants.

In a later study, Christensen et al. (2018) investigated WM impairments in aphasia. The study involved 14 individuals with aphasia and compared their performance with 13 control participants. To measure WM impairments, the study employed verbal and spatial *n*-back tasks, the Flanker test, as well as forward digit and spatial tasks. For the verbal task, the study employed the verbal *n*-back task, which included letters. For the spatial task, the study used an *n*-back task with black circles. Christensen et al.'s (2018) study revealed that aphasic participants demonstrated poorer performance in the verbal tasks compared to the control group. However, in spatial tasks, both groups performed similarly, suggesting that aphasics tend to exhibit significant challenges in WM in the context of verbal tasks. The study also indicated that impairments in verbal WM in aphasics could be due to inhibition problems, e.g., an inability to suppress irrelevant information, which might also play a role in working memory deficits.

In addition to investigating non-linguistic working memory in individuals with aphasia, recent studies have focused on the efficacy of WM training in aphasia. For example, Zakariás et al. (2018) investigated whether the benefits of computerised WM training would help to improve not only the participants' working memory skills but also auditory comprehension. The training targeted various components of working memory. The study involved three individuals with aphasia in the chronic phase of recovery. The results showed that aphasics demonstrated some improvement regarding at least one measure of auditory comprehension and everyday memory tasks. Moreover, two aphasics exhibited some improvement in measures of working memory and functional communication skills. Zakariás et al.'s (2018) study revealed that the WM performance could be enhanced by a computer-based training in the chronic stage of recovery in individuals diagnosed with aphasia.

Furthermore, Nikravesch et al. (2021) investigated how WM training affects people with aphasia in memory and language performance. The study involved 13 participants diagnosed with aphasia with moderate and mild symptoms. Their findings revealed that individuals with aphasia exhibited notable improvements in memory tasks. Additionally, the language performance of individuals with aphasia improved as well, suggesting that working memory training could be beneficial not only for memory but also for language abilities in individuals with aphasia.

1.4. Working memory in bilingual aphasia

Data on the updating component of WM in bilingual aphasics are limited (Penn et al., 2009; Adrover-Roig et al., 2011; Lee et al., 2016; Penn et al., 2017) and of mixed outcome.

The preliminary study of Penn et al. (2009) investigated impairments in updating in bilinguals with aphasia in the chronic stage of recovery. The study involved two bilingual and eight monolingual aphasics with mild and moderate symptoms. The Working Memory Self-Ordered Pointing Test used by the authors did not reveal impairments in the two bilingual patients, whereas impaired updating was demonstrated in the eight monolingual aphasics.

Furthermore, Adrover-Roig et al. (2011) have investigated a 53-year-old right-handed Basque-Spanish bilingual male who had a haemorrhagic incident of stroke. The patient, at the time of investigation, was in the chronic stage of his recovery and showed selective impairment, e.g., his L2 was more intelligible than his L1, even after 3 months of speech therapy in both languages. The brain damage in the patient was restricted to the left basal ganglia, resulting in impairments in his L1 to a larger extent. The language assessment tools included the Bilingual Aphasia Test and the Boston Naming Test in both languages. The executive function assessments included the forward and backward digit span tasks, the classical Stroop test, and the Trail Making Test (TMT). The results of these

executive function assessments revealed impairments in updating, inhibition, and shifting.

Lee et al. (2016) described the impairments of a bilingual patient with crossed aphasia, a rare case in dextrals, e.g., a right hemispheric lesion in a right-handed person. The 47-year-old male patient had an extensive haemorrhage in the right basal ganglia. The investigation utilized the Western Aphasia Battery (WAB), which was administered in the subacute stage of recovery and later in the chronic stage, following a six-month course. This resulted in balanced proficiency in both languages, with equal abilities in the L1 (Korean) and L2 (Japanese). The WAB was employed in both languages, with AQ scores of 32.0 for Korean and 50.6 for Japanese, respectively. Based on the WAB results, the patient was diagnosed with Broca's aphasia. The fMRI (functional Magnetic Resonance Imaging) study revealed left lateralised activation in both language tasks, especially in the inferior frontal gyrus. The non-linguistic assessments were conducted in the subacute stage using a range of tests, including the digit span task, the TMT, the Rey Complex Figure Test, and the Wisconsin Card Sorting Test (WCST). The results of these tests revealed multiple impairments in the subacute stage of recovery, including inhibition, updating, and shifting. Lee et al. (2016) revealed that, despite the absence of a follow-up assessment of the executive functions in the chronic stage, repetition of the WAB and the naming test showed no sign of the paradoxical pattern of bilingual aphasia.

Finally, Penn et al. (2017) investigated the recovery patterns during the acute and subacute stages of recovery in post-stroke participants with bilingual aphasia (N = 10). The ten aphasic patients (seven males and three females aged 27-65 years) with different sites of and types of aphasia were tested over two periods of time within the first 12 weeks following the incident leading to severe symptoms of aphasia. The control group consisted of 19 neurologically intact bilingual individuals (6 males and 13 females) and was matched as closely as possible to the aphasic patients in terms of age, level of education, and linguistic variety. Penn et al.'s (2017) study revealed impaired updating, as tested by a non-linguistic n-back task, in individuals with aphasia. It is, however, noteworthy that they exhibited impairment both in the acute and subacute stages of recovery.

1.5. The present study

The present study aimed to investigate working memory impairments in individuals with monolingual and bilingual aphasia. Therefore, visual and auditory working memory tasks were employed.

Our first hypothesis was that participants with aphasia would demonstrate impaired working memory in the visual domains. To identify the expected impairments, the Visual Sequential Memory Test and the Rey-Osterrieth Complex Figure B Test were employed.

Our second hypothesis was that verbal working memory impairments would be identified by using the Auditory Sequential Memory Test.

Our third hypothesis was that bilinguals with aphasia would perform better in working memory tests than monolingual aphasics.

Our fourth hypothesis was that significant performance differences would be found between aphasic participants and neurologically healthy participants, serving as a matched reference, for both bilinguals and monolinguals.

2. Methods

2.1. Participants

The study aimed to investigate working memory capabilities of individuals with monolingual and bilingual aphasia and compare with neurologically intact mono- and bilingual participants as a reference. A total of 54 participants were recruited for this study, divided into four groups: monolingual people with aphasia (mPWA), bilingual people with aphasia (bPWA), monolingual healthy participants (mHP), and bilingual healthy participants (bHP), all of whom had no neurological diagnosis. All bilingual participants have used the two languages in parallel. All participants provided informed consent, and the confidentiality of their data was maintained through the anonymization process, which involved assigning a unique identification number to each participant. Participants were informed of their right to withdraw from the study at any time without consequences. Tables 2 and 3 present the basic demographic data (age, sex, years of education) and clinical data of aphasic patients. Tables 4 and 5 present the basic data for the reference groups. Data on L1 and L2, as well as the age of acquisition, are presented in Tables 3 and 5.

2.1.1. Aphasia groups

All participants with aphasia had suffered a cerebrovascular accident. Clinical criteria for aphasic participants included: (1) signs of aphasia as indicated by their performance during the WAB test, (2) at least 6 months post-onset of their stroke, and (3) premorbid right-handedness.

Table 2. Demographic data of the monolingual aphasia group

Participants	Age	Sex	Education in years	Post-onset days	Speech	Lesion site	Aphasia severity
mPWA1	49	female	11	395	non-fluent	left	severe
mPWA2	76	male	11	556	fluent	left	mild
mPWA3	71	male	11	181	fluent	right and left	mild
mPWA4	63	male	11	183	non-fluent	left	severe
mPWA5	71	male	11	10717	fluent	left	mild
mPWA6	69	female	16	6162	fluent	left	mild
mPWA7	61	male	11	2363	non-fluent	left	severe
mPWA8	83	female	11	278	fluent	left	mild
mPWA9	64	male	22	730	fluent	left	mild
mPWA10	80	female	12	183	fluent	left	moderate
Mean	68.7		12.7	2174.8			
SD	9.99		3.62	3532.16			
min-max	49-83		11-22	181-10717			

The monolingual aphasia group consisted of 10 participants. The demographic data for this group is summarised in Table 2. Participants ranged in age from 49 to 83 years, with a mean age of 68.7 years ($SD = 9.99$). The group comprised six males and four females, all of whom had completed 11 to 22 years of education, with a mean education level of 12.7 years ($SD = 3.62$). The duration of post-onset days varied among participants, ranging from 181 to 10,717 days, with a mean of 2174.8 days ($SD = 3532.16$). Speech characteristics varied, with some participants exhibiting fluent speech ($N = 7$) and others classified as non-fluent ($N = 3$). Lesion sites were predominantly located in the left hemisphere, and aphasia severity ranged from severe to mild.

Table 3. Demographic data of the bilingual aphasia group

Participants	Age	Sex	Education in years	Post-onset days	Speech	Lesion site	Aphasia severity (parallel)	L1	AoA L1	L2	AoA L2
bPWA1	73	female	12	8580	fluent	left	mild	CRO	from birth	HU	6
bPWA2	72	female	16	603	fluent	left	mild	CRO	from birth	HU	6
bPWA3	57	male	12	8978	fluent	left	mild	CRO	from birth	HU	3
bPWA4	60	male	11	369	fluent	left	mild	HU	from birth	DE	30
bPWA5	49	female	20	225	fluent	left	mild	HU	from birth	SK	3
bPWA6	72	male	11	201	non-fluent	left	severe	HU	from birth	SK	6
bPWA7	44	female	14	1989	fluent	left	moderate	HU	from birth	SK	3
bPWA8	81	female	11	184	non-fluent	left	severe	HU	from birth	SK	6
Mean	63.5		13.38	2641.13							7.88
SD	13.02		3.20	3835.87							9.06
min-max	44-81		11-20	184-8978							3-30

The bilingual aphasia group consisted of 8 participants, with demographic details presented in Table 3. Participants' ages ranged from 44 to 81 years, with a mean age of 63.5 years ($SD = 13.02$). This group included three males and five females, with educational backgrounds ranging from 11 to 20 years (mean = 13.38 years, $SD = 3.20$). The post-onset days varied from 184 to 8978 days, with a mean of 2641.13 days ($SD = 3835.87$). Six participants reported fluency in their speech, and two individuals were classified as non-fluent. The participants had Croatian (CRO) or Hungarian (HU) as their first language (L1) and had acquired their second language (L2), which was either Hungarian, Slovakian (SK), or German (DE), at varying ages.

2.1.2. Reference groups

The healthy monolingual group consisted of 18 participants, with their demographic data summarised in Table 4. Participants' ages ranged from 45 to 85 years, with a mean age of 67.67 years ($SD = 10.99$), which is highly comparable with the aphasia group. This group consisted of 8 males and 10 females, with education levels ranging from 8 to 22 years (mean = 13.28 years, $SD = 3.63$). This group served as a reference for the monolingual aphasia group.

Table 4. Demographic data of the healthy monolingual participants

Participants	Age	Sex	Education in years
mHP1	70	female	14
mHP2	71	male	11
mHP3	72	female	16
mHP4	75	female	8
mHP5	66	male	11
mHP6	64	female	11
mHP7	63	female	15
mHP8	63	male	11
mHP9	65	female	12
mHP10	45	female	16.5
mHP11	85	male	8
mHP12	64	male	14
mHP13	48	female	16
mHP14	58	male	15.5
mHP15	75	male	14
mHP16	85	female	8
mHP17	83	female	16
mHP18	66	male	22
Mean	67.67		13.28
SD	10.99		3.63
min -max	45-85		8-22

The healthy bilingual group consisted of 18 participants, whose demographic characteristics are presented in Table 5. Participants' ages ranged from 46 to 80 years, with a mean age of 65 years ($SD = 9.80$). The group consisted of 13 females and five males, with educational attainment ranging from 8 to 30 years (mean = 13.17 years, $SD = 5.51$). Similarly to the bilingual aphasia group, the participants had their first language (L1) as Croatian (CRO) or Hungarian (HU) and had acquired their second language (L2), which was either Hungarian, Croatian, Slovakian (SK), or German (DE), at varying ages. This group served as a reference for the bilingual aphasia group.

Table 5. Demographic data of the healthy bilingual participants

Participants	Age	Sex	Education in years	L1	AoA L1	L2	AoA L2
bHP1	62	female	12	CRO	from birth	HU	3
bHP2	63	female	11	CRO	from birth	HU	3
bHP3	66	female	16	CRO	from birth	HU	3
bHP4	66	female	14	CRO	from birth	HU	6
bHP5	67	female	8	CRO	from birth	HU	6
bHP6	67	female	16	CRO	from birth	HU	6
bHP7	71	female	8	CRO	from birth	HU	3
bHP8	72	male	8	CRO	from birth	HU	3
bHP9	72	male	11	CRO	from birth	HU	6
bHP10	76	female	8	CRO	from birth	HU	6
bHP11	77	female	11	CRO	from birth	HU	6
bHP12	69	male	11	HU	from birth	DE	30
bHP13	50	female	18	HU	from birth	SK	from birth
bHP14	80	female	8	CRO	from birth	HU	6
bHP15	60	female	30	HU	from birth	CRO	from birth
bHP16	46	female	18	HU	from birth	CRO	from birth
bHP17	47	male	17	CRO	from birth	HU	3
bHP18	59	male	12	HU	from birth	CRO	23
Mean	65		13.17				6.28
SD	9.80		5.51				7.77
min-max	46-80		8-30				0-30

2.2. Procedures

Participants with aphasia were recruited from various sources. The recruitment was conducted via three principal channels: (a) institutional, (b) private speech and language therapists, and (c) individuals.

- (a) Institutional recruitment: Some participants were recruited from rehabilitation centres and hospitals, approached in collaboration with rehabilitation departments where speech therapists identified suitable candidates based on predefined inclusion criteria.

The individuals bPWA2, mPWA2, and mPWA7 were recruited from the Sopron Erzsébet Teaching Hospital and Rehabilitation Institute in Hungary (Soproni Erzsébet Oktató Kórház és Rehabilitációs Intézet). The hospital provided a speech therapist, a bilingual psychologist for the Croatian tests, and a quiet, private room for the assessments. The speech therapist conducted the Hungarian tests with bPWA2, mPWA2, and mPWA7, while the bilingual psychologist carried out the Croatian tests with bPWA2. Both the Hungarian and Croatian assessments were conducted in the presence of the author of this study. The hospital director of the Rehabilitation Institute in Sopron granted written permission (reference number: 543-2/2023) for the first author of this study to enter the Institution and utilise the speech therapy room to conduct assessments.

Participant bPWA5 was recruited from Penta Hospitals (Nemocnica Dunajská Streda) in Slovakia. The director of the hospital gave written authorisation to the first author of this study to enter the Institution and use the speech therapy room to conduct assessments. The hospital provided a speech therapist and a suitable room for the assessments. Since the speech therapist was bilingual (L1 Hungarian, L2 Slovakian), all the tests were conducted in the presence of the speech therapist and the author of this study.

- (b) Private Speech Therapists: Additional participants were recruited through private speech therapists. These professionals were contacted to identify individuals who are actively seeking therapy and meet the study's eligibility requirements. Written informed consent was obtained from both therapists and participants prior to collecting data. Individuals bPWA6, bPWA7, and bPWA8 were recruited from a private speech therapist in Slovakia, while individuals mPWA8, mPWA9, and mPWA10 were recruited from a private speech therapist in Hungary.

- (c) Individual recruitment: A subset of participants was recruited on an individual basis, particularly those who had completed a course of formal speech therapy and were no longer receiving regular treatment. These individuals were approached directly, often through community outreach or support groups for people with aphasia. Data for this group was collected in their homes to ensure a comfortable and familiar environment, thereby

facilitating a more effective assessment. The following participants were recruited individually: bPWA1, bPWA3, bPWA4, mPWA1, mPWA3, mPWA4, mPWA5, and mPWA6. All participants provided written informed consent.

Due to the length of the tests, the examinations were performed in several sessions. Every participant was tested individually. All procedures were conducted in accordance with ethical guidelines, ensuring that the participants were fully informed of the study's purpose and their right to withdraw their participation at any time.

2.3. Data analysis

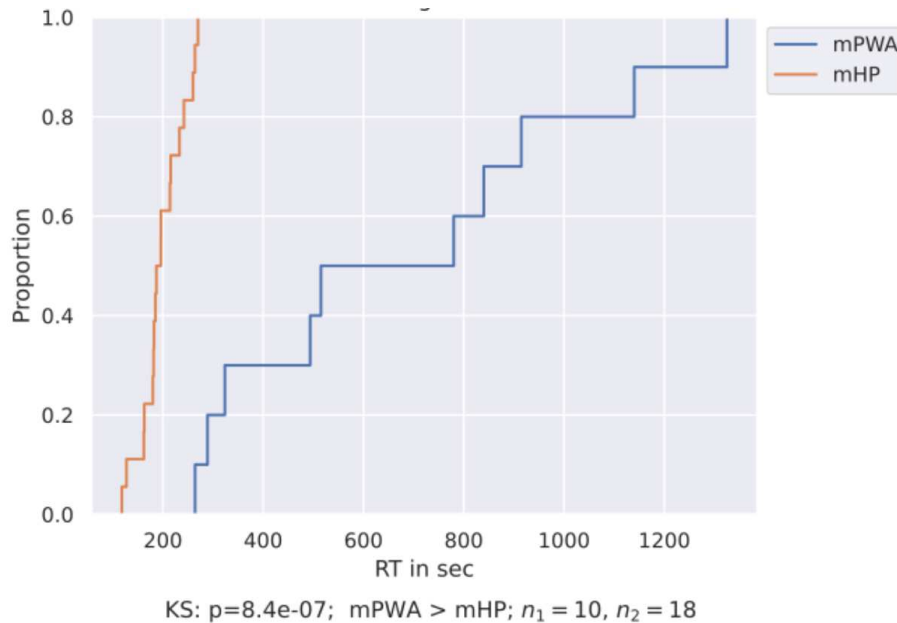
The Kolmogorov-Smirnov (KS) test, a non-parametric and distribution-free test that imposes no assumptions regarding the underlying data distribution, was used. Specifically, the two-sample KS test with a one-sided alternative hypothesis for stochastic dominance was employed. Unlike the Mann–Whitney U test, which primarily evaluates differences in central tendencies between two populations, the KS test considers the entire distribution of data (Dodge, 2008). In the case of discrete data, the KS test tends to be conservative, as noted in previous studies (Noether, 1963; Slakter, 1965; Walsh, 1963). We conducted our analysis using the SciPy library's implementation (Virtanen et al., 2020), specifically with the `kstest` function.

The empirical cumulative distribution function (ECDF) is a useful statistical tool for visualising the distribution of data points in a dataset. In our study, ECDF plots were used to illustrate the differences in performance between individuals with bilingual and monolingual aphasia as well as between individuals with aphasia and healthy participants.

Figure 1 illustrates an example of an ECDF graph, where the x-axis represents the values of the variable being studied, e.g., response time or test scores. The y-axis represents the cumulative proportion of participants scoring at or below a given score. For instance, a point on the stepped line at 0.6 indicates 60% of the observations fall below this value. The line rises in steps, with each step corresponding to an observation in the dataset. In our example (Figure 1), we compare two groups: individuals with monolingual aphasia (mPWA) and healthy monolingual participants (mHP) in a task that assesses performance differences in response time between the two groups. It can be seen in the example that the $p\text{-value} \text{ and } mPWA > mHP$ are included in the figure. The relation sign means that the mPWA group produced greater numerical data compared to the mHP group. It is important to note that the stepped line of the group with the greater numerical data is constantly situated below the other line. The steepness of the lines can also provide insight into the variability of performance within each group. In our

example, a steep stepped line (orange for mHP) indicates less variability, while a flatter stepped line (blue for mPWA) suggests more variability in performance.

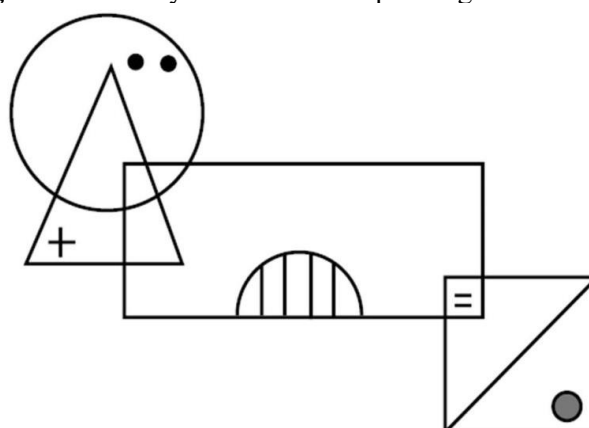
Figure 1. An example of an ECDF graph.



2.4. Assessment tools

The Rey-Osterrieth Complex Figure B Test (ROCF-B) is a valuable tool for assessing cognitive impairment, especially in individuals with aphasia (Lezak et al., 2012). The test assesses a range of cognitive functions, including visuospatial skills, visual memory, and executive functions, by requiring participants to copy and recall a complex geometric figure. The copying phase primarily assesses visuospatial skills, perceptual organization, and motor planning. After a three-minute delay, the participants were required to draw the figure from memory. This recalling phase assessed visual memory, organizational strategies, planning, monitoring, and inhibitory control (Kirkwood et al., 2001). Figure 2 illustrates the drawing included in the ROCF-B test. The maximum achievable score was 31. The scoring instructions included the following:

- The number of drawn elements (11 points): a circle, two small points, a triangle, a cross, a rectangle, an arc, a square, a diagonal line, an equal sign, a large point, four lines.
- The position of elements (8 points): a cross in the triangle, two points in the circle, a large point in the square, an arc in the rectangle, four lines inside the arc, the square, the equal sign, the diagonal line in the square.
- Overlaps (8 points): circle-triangle, triangle-rectangle, circle-rectangle, rectangle-square.
- The proportions between the four principal elements (4 points).

Figure 2. The Rey-Osterrieth Complex Figure B Test

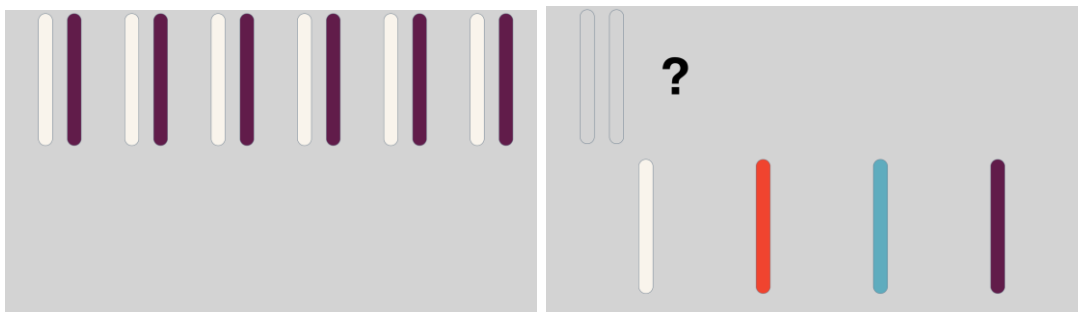
In the Auditory Sequential Memory Test, the participants were required to match words to corresponding pictures (see Figure 3). A trial test was conducted to determine if the participants understood and could carry out the task. If successful, a sequence of unrelated words was read aloud, and the participants were asked to reproduce the sequence by pointing to the corresponding pictures. The test began with a sequence of two words, which gradually increased to a maximum of nine. The maximum score achievable was 44. The images were sourced from a Hungarian Aphasia Therapy Exercise Book (Szabó & Sándor, 2021). The task measured working memory as well as cognitive flexibility and inhibition.

Figure 3. Pictures included in the Auditory Sequential Memory Test

A Visual Sequential Memory Test was applied from a Hungarian rehabilitation software which was made specifically for post-stroke patients (called Afázia App created by Katalin Nádudvari and Ali Baharev, PhD, available online at <https://afazia.app/>), involving coloured sticks and requiring the participants to memorise the order of 12 such sticks (Cseh & Hegyi, 1995, pp. 7-27). Each part of the test involved two colours, such as white and purple, but these had to be

chosen from four different colours, namely white, red, blue, and purple (refer to Figure 4). The software instantly provided feedback to the participant, indicating whether the selected colour was correct or incorrect. All participants completed the test on the same tablet to ensure a consistent assessment environment. Their response times were measured in seconds to investigate the differences in performance between individuals with aphasia and those without. The test evaluated visual working memory and cognitive flexibility.

Figure 4. An example of the Visual Sequential Memory Test



3. Results

This section presents the results of the study investigating working memory impairments in participants with aphasia, specifically focusing on both the visual and verbal domains. The analysis involved two groups of aphasics: bilingual people with aphasia (bPWA) and monolingual people with aphasia (mPWA), as well as two reference groups: healthy bilingual participants (bHP) and healthy monolingual participants (mHP). The results are organised according to the three tests employed: the Visual Sequential Memory Test (VSMT), the Rey-Osterrieth Complex Figure B Test (ROCF-B), and the Auditory Sequential Memory Test (ASMT). The results are discussed in terms of response times (RT), percentages of mistakes, and overall performance scores, highlighting the implications of cognitive load for bilingual individuals compared to their monolingual counterparts. The individual scores of the bilingual and monolingual groups are displayed in Tables 6 and 7. As illustrated in Table 8, the results of the statistical analysis are presented alongside their respective p-values (*significant with Italic*).

Table 6. Results of the auditory and visual working memory tasks in the group of bilingual aphasics and the bilingual reference group.

Participants		Visual sequential memory test (mean RT in sec.)	Visual sequential memory (mistakes pct.)	ROCF-B (max score 31)	Auditory sequential memory test (max score 44)
Group of bilingual aphasics (N=8)	bPWA1	3.56	9.09	26	35
	bPWA2	1.34	2.70	30	36
	bPWA3	1.69	5.88	12	18
	bPWA4	2.19	7.69	27	44
	bPWA5	1.18	0	30	43
	bPWA6	3.55	22.58	0	0
	bPWA7	1.36	0	27	35
	bPWA8	5.74	40.00	0	3
Range of scores (min-max)		1.18-5.74	0-40	0-30	0-44
Mean score		2.58	10.99	19	26.75
SD		1.60	13.77	13.04	17.48
Bilingual reference group (N=18)	bHP1	1.19	0	21	42
	bHP2	1.11	2.04	31	44
	bHP3	1.24	0	21	44
	bHP4	1.12	0	31	44
	bHP5	1.28	4.00	26	35
	bHP6	1.44	2.04	31	42
	bHP7	1.55	5.88	15	29
	bHP8	1.19	4.00	31	35
	bHP9	2.03	4.00	26	36
	bHP10	1.38	7.69	25	38
	bHP11	1.8	0	27	34
	bHP12	1.49	4.00	23	42
	bHP13	1.02	0	31	44
	bHP14	2.04	0	22	35
	bHP15	1.06	0	26	42
	bHP16	1.17	0	24	42
	bHP17	1.07	0	29	44
	bHP18	1.01	0	26	40
Range of scores (min-max)		1.01-2.04	0-7.69	15-31	29-44
Mean score		1.34	1.87	25.89	39.56
SD		0.33	2.46	4.44	4.53

Table 7. Results of the auditory and visual working memory tasks in the group of monolingual aphasics and the monolingual reference group.

Participants		Visual sequential memory test (mean RT in sec.)	Visual sequential memory test (mistakes pct.)	ROCF-B (max score 31)	Auditory sequential memory test (max score 44)
Group of monolingual aphasics (N=10)	mPWA1	2.11	7.69	22	42
	mPWA2	3.2	7.69	11	30
	mPWA3	1.94	0	24	42
	mPWA4	4.85	45.45	13	23
	mPWA5	1.37	0	31	43
	mPWA6	1.21	4.35	24	20
	mPWA7	5.22	20.00	21	16
	mPWA8	3.09	14.29	17	26
	mPWA9	1.20	0	29	44
	mPWA10	2.49	11.11	7	14
Range of scores (min-max)		1.20-5.22	0-45.45	7-31	14-44
Mean score		2.67	11.06	19.90	30
SD		1.44	13.77	7.80	11.88
Monolingual reference group (N=18)	mHP1	1.27	0	31	44
	mHP2	1.05	0	30	44
	mHP3	1.21	0	15	44
	mHP4	2.27	4.00	30	37
	mHP5	1.23	2.04	31	44
	mHP6	1.31	0	28	39
	mHP7	1.15	0	24	40
	mHP8	2.04	6.25	22	30
	mHP9	1.47	0	31	44
	mHP10	1.16	0	29	44
	mHP11	1.82	17.24	23	42
	mHP12	2.05	5.26	22	41
	mHP13	1.04	0	27	44
	mHP14	1.28	5.26	25	40
	mHP15	2.56	7.69	31	42
	mHP16	1.88	14.29	15	26
	mHP17	0.99	0	31	42
	mHP18	1.43	0	24	40
Range of scores (min-max)		0.99-2.56	0-17.24	15-31	26-44
Mean score		1.51	3.45	26.06	40.39
SD		0.48	5.21	5.22	5.02

3.1. Results of the Visual Sequential Memory Test

3.1.1. Monolingual aphasics vs. bilingual aphasics

Results from the Visual Sequential Memory Test, administered to the bilingual aphasic participants, indicated a mean response time (RT) of 2.58 seconds (sec), with a standard deviation (SD) of 1.60. The average percent of mistakes made by this group was 10.99%, with an SD of 13.77. The range of errors varied from 0 to 40%, indicating a high degree of variance in performance among the participants. Individual levels of performance varied considerably; for instance, bPWA8 recorded the highest mean RT of 5.74 sec, and 40 % of the answers were wrong, while bPWA5 demonstrated the most efficient level of performance with a mean RT of 1.18 sec and no mistakes. This suggests that increased response times may correlate with a higher incidence of errors within the aphasia group.

In contrast, the mean RT of the monolingual aphasia group was 2.67 sec (SD = 1.44) in the VSMT, with an average percentage of mistakes made of 11.06% (SD = 13.77). The response times ranged from 1.20 to 5.22 seconds, exhibiting lower variability in performance compared to the bilingual aphasia group. The percentage of mistakes made for the monolingual aphasia group varied from 0 to 45.45%, once again highlighting the variability of performance. For example, the RTs and percentages of mistakes made for mPWA4 and mPWA7 were high, while the RTs for mPWA3 and mPWA9 were low, with no mistakes made. This suggests that, similar to the bilingual aphasia group, there may be a correlation between increased response times and the number of errors made.

In conclusion, although the mean RT for the bilingual aphasia group was lower than that for the monolingual aphasia group, the differences revealed a trend that did not reach statistical significance, and the same was valid for the percentage of mistakes. In both groups, participants with more severe aphasia symptoms demonstrated poorer performance than those with moderate or mild aphasia symptoms. As we hypothesised, the response time increased, and the number of mistakes rose, showing a positive correlation between response time and percentage of errors made for both aphasia groups.

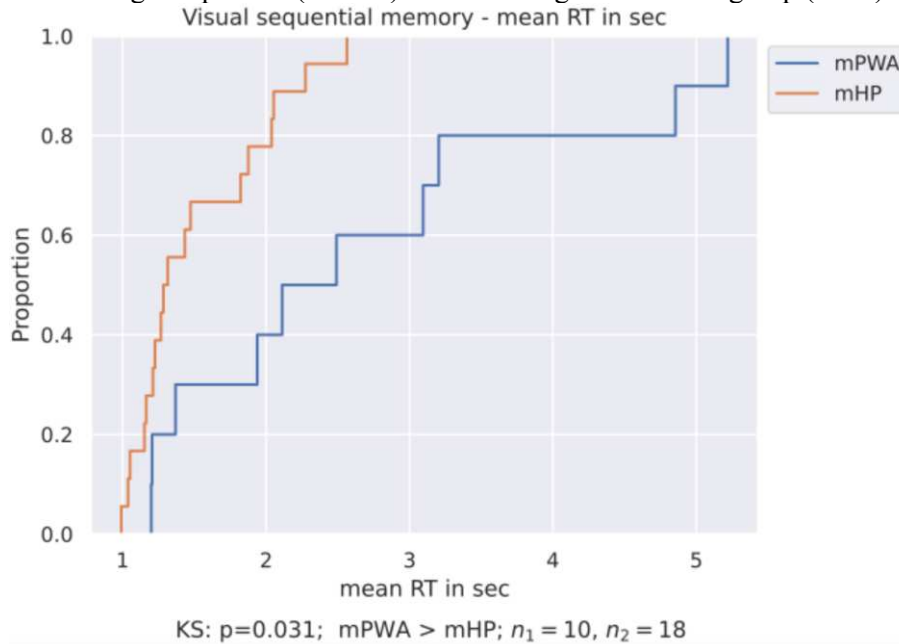
Table 8. Results from the statistical analysis of auditory and visual working memory tasks.

	Auditory sequential memory number of correct answers	ROCF-B score	Visual sequential memory	
			mean RT in sec.	mistakes pct.
mPWA (N=10) vs. bPWA (N=8)	p=0.514	p=0.148	p=0.623	p=0.737
mPWA (N=10) vs. mHP(N=18)	<i>p = 0.031</i> <i>aphasia <</i> <i>reference</i>	p=0.084	<i>p=0.031</i> <i>aphasia ></i> <i>reference</i>	p=0.066
bPWA(N= 8) vs. bHP (N=18)	p > 0.10	p > 0.10	<i>p=0.044</i> <i>aphasia ></i> <i>reference</i>	<i>p=0.036</i> <i>aphasia ></i> <i>reference</i>

3.1.2. Monolingual aphasics vs. monolingual reference group

The mean response time (RT) for the Visual Sequential Memory Test was significantly longer in the monolingual aphasia group in comparison with the monolingual reference group (refer to Figure 5). While the monolingual aphasia group exhibited a mean RT of 2.67 sec (SD = 1.44), the monolingual reference group performed the task with a mean RT of 1.51 sec (SD = 0.48). This difference indicates a notable impairment in visual sequential memory among individuals with aphasia. The range of RT scores for the monolingual aphasia group spanned from 1.20 to 5.22 sec, whereas the reference group's RT scores ranged from 0.99 to 2.56 sec. Additionally, we also assessed the percentage of mistakes made during the Visual Sequential Memory Test. The mean percentage of mistakes for the monolingual aphasia group was 11.06% (SD = 13.77), and for the reference group was 3.45% (SD = 5.21), without a statistically significant difference.

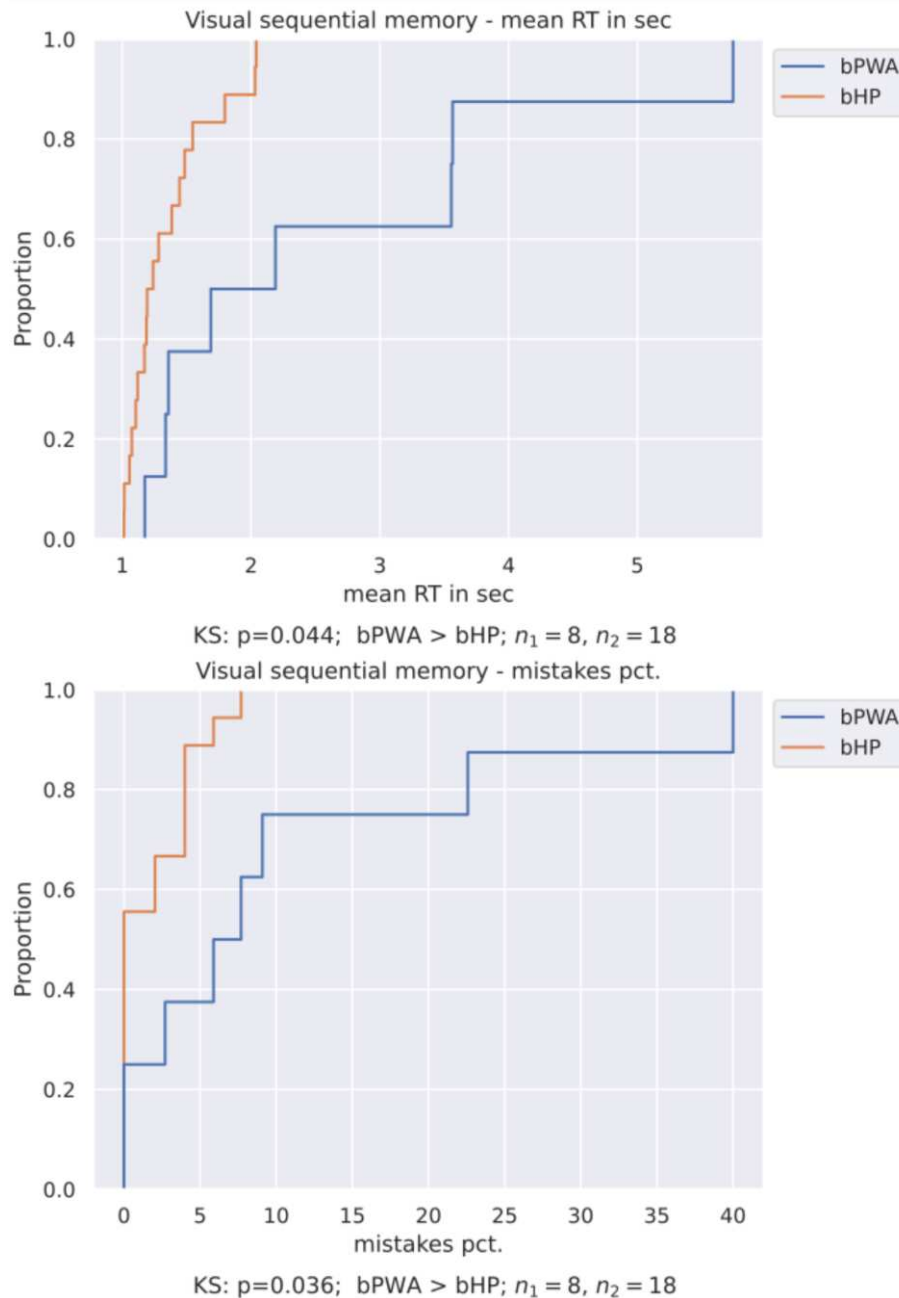
Figure 5. Comparison of the results from the Visual Sequential Memory Test recorded by the monolingual aphasics (mPWA) and monolingual reference group (mHP).



3.1.3. Bilingual aphasics vs. bilingual reference group

The response time (RT) results of the Visual Sequential Memory Test indicated a considerable degree of variability within the bilingual aphasia group. The mean RT for this group was 2.58 seconds ($SD = 1.60$), with a range from 1.18 to 5.74 seconds. In contrast, the bilingual reference group ($N=18$) demonstrated a significantly lower mean RT of 1.34 seconds ($SD = 0.33$) with scores ranging from 1.01 to 2.04 seconds. In addition, the bilingual aphasia group had a mean error rate of 10.99% ($SD = 13.77$) with a range of 0% to 40%. Conversely, the bilingual reference group exhibited a mean error rate of only 1.87% ($SD = 2.46$), with scores ranging from 0 to 7.69%. The higher error rates in the bilingual aphasia group suggest a well-expressed impairment of the visual working memory and reveal a double deficit in terms of response time and accuracy. This finding aligns with our first hypothesis that participants with aphasia underperform in visual memory tasks. As shown in Figure 6, a statistically significant difference was found between the two groups in both response time and the percentage of mistakes.

Figure 6. Comparison of the results from the Visual Sequential Memory Test recorded by the bilingual aphasics (bPWA) and the bilingual reference group (bHP).



3.2. Results from the Rey Complex Figure-B

3.2.1. Monolingual aphasics vs. bilingual aphasics

The results of the ROCF-B further elucidated the differences in cognitive processes assumed and observed in the tasks described above. The performance of the bilingual aphasia group in the ROCF-B test yielded a mean score of 19 (SD = 13.04), with scores ranging from 0 to 30, indicating a moderate level of impairment in visual memory as assessed by this test. While bPWA1 achieved a score of 26, bPWA6 and bPWA8, with more severe aphasia symptoms, scored the lowest, that is, 0. The monolingual aphasia group produced a mean score of 19.90

(SD = 7.80) in the ROCF-B, with scores ranging from 7 to 31. The bilingual individuals with aphasia scored lower on average than their monolingual aphasic counterparts, with no statistical difference.

3.2.2. Monolingual aphasics vs. monolingual reference group

The monolingual aphasia group had a mean score of 19.90 (SD = 7.80) out of a maximum score of 31, while the monolingual reference group achieved a mean score of 26.06 (SD = 5.22). This difference in mean scores suggests that individuals with aphasia demonstrate deficits in visual working memory compared to their healthy counterparts, although this trend is only apparent.

3.2.3. Bilingual aphasics vs. bilingual reference group

The bilingual aphasia group achieved a mean score of 19 (SD = 13.04), with scores ranging from 0 to 30. In contrast, the bilingual reference group had a mean score of 25.89 (SD = 4.44), with scores ranging from 15 to 31. Within the bilingual aphasia group, those exhibiting more severe aphasia symptoms demonstrated poorer performance compared to their counterparts (see Table 6 for individual scores). However, no statistical significance was found between the two groups.

3.3. Results from the Auditory Sequential Memory Test

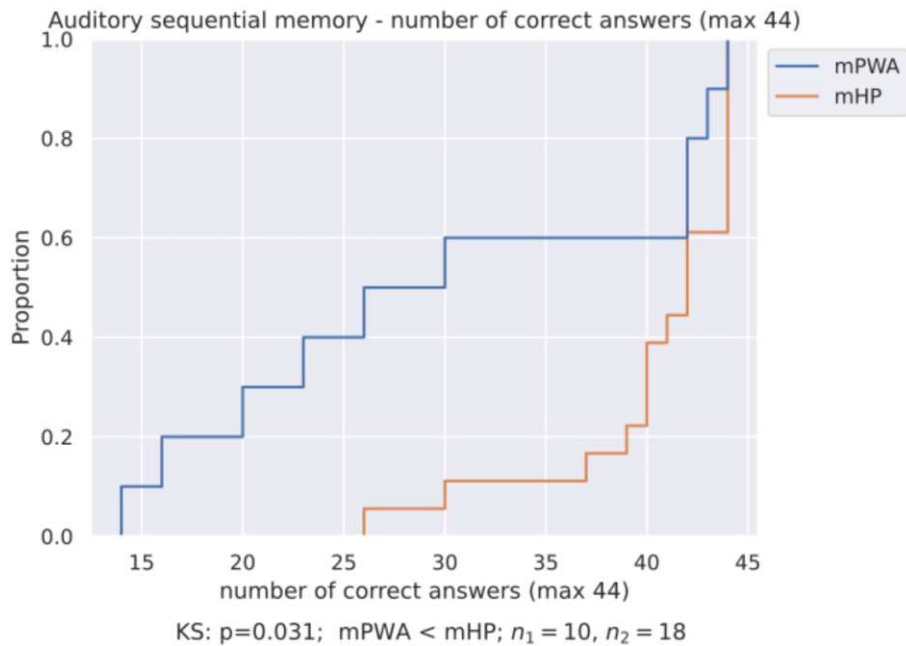
3.3.1. Monolingual aphasics vs. bilingual aphasics

The results from the Auditory Sequential Memory Test produced a mean score of 26.75 (SD = 17.48) for the bilingual aphasic participants, with scores ranging from 0 to 44, indicating a substantial level of impairment in verbal working memory. The monolingual aphasic participants achieved a mean score of 30 (SD = 11.88) in the ASMT, with scores ranging from 14 to 44. The analysis revealed that the monolingual aphasia group performed better than the bilingual aphasia group in the Verbal Working Memory Test, a trend that was not statistically significant.

3.3.2. Monolingual aphasics vs. monolingual reference group

The Auditory Sequential Memory Test results also revealed significant differences between the monolingual aphasia and reference groups (see Figure 7). The monolingual aphasia group exhibited a mean score of 30 (SD = 11.88) out of a maximum score of 44, while the reference group produced a mean score of 40.39 (SD = 5.02). This difference suggests that individuals with aphasia also experience impairments in auditory sequential memory. The scores recorded by the monolingual aphasia group ranged from 14 to 44, demonstrating variability in performance, with some individuals achieving scores close to the maximum, while others scored considerably lower. In contrast, the scores from the monolingual reference group ranged from 26 to 44, indicating the consistently better level of performance of these participants.

Figure 7. Comparison of the results from the Auditory Sequential Memory Test between the monolingual aphasics (mPWA) and the monolingual reference group (mHP).



3.3.3. Bilingual aphasics vs. bilingual reference group

To assess verbal working memory, we employed the Auditory Sequential Memory Test. The bilingual aphasia group had a mean score of 26.75 (SD = 17.48), with scores ranging from 0 to 44. In comparison, the bilingual reference group achieved a higher mean score of 39.56 (SD = 4.53), with scores ranging from 29 to 44. The findings suggest a potential deficit in auditory working memory among bilingual individuals with aphasia, although this is only a trend.

4. Discussion

The present study aimed to investigate working memory impairments in individuals with aphasia, with a specific focus on differences between bilingual and monolingual individuals with aphasia by using tests for the visual and auditory domains. The findings from the Visual Sequential Memory Test, the Rey Complex Figure Test, and the Auditory Sequential Memory Test provided critical insights into the cognitive load and performance discrepancies experienced by these aphasia groups.

Statistical analyses of results found in the investigated monolingual and bilingual aphasia groups did not reveal any significant differences, neither in the visual nor in the auditory working memory tests. The results of the Visual Sequential Memory Test confirmed significant performance variability in both the mono- and bilingual aphasia groups, as indicated by the wide range of response times and error rates. The mean response time difference between the two aphasia groups showed trends only and not statistical significance, suggesting similarities in cognitive load when engaged in visual memory tasks. The variability in

performance observed in the aphasia groups, including participants who demonstrated markedly higher response times and error rates, indicates that individual differences might play a crucial role in the cognitive performance of these patients. Both the mono- and bilingual aphasia groups demonstrated impairments in visual memory as shown by the Rey Complex Figure B (ROCF-B) test. Again, the bilingual aphasia group achieved lower mean scores than the monolingual aphasia group, although the difference was not statistically significant. The variability in scores observed within the bilingual aphasia group, particularly among those with more severe aphasia symptoms, highlights the need for tailored interventions that focus on the clinical aspects of individual differences. Moreover, the Auditory Sequential Memory Test revealed a substantial level of impairment in verbal working memory both in the mono- and bilingual aphasia groups, with bilinguals scoring lower than monolinguals. Again, a trend is evident, but it lacks statistical significance. As expected, the aphasia and reference groups showed statistically significant differences both in the visual and auditory working memory tests.

Our finding of a significant impairment in visual sequential memory among individuals with aphasia is consistent with the existing literature, which highlights the complexity of cognitive deficits associated with aphasia, particularly in memory tasks (Mooijman et al., 2022). However, there were several tasks, such as the ROCF-B test, where the performance difference between aphasics and healthy participants showed trends similar to those in the Rey Complex Figure B test, but without statistical significance. Therefore, a logical question is whether these weaknesses add up in complex tasks and how they contribute to performance dissociations.

The results from the Auditory Sequential Memory Test (ASMT) further emphasise the role of auditory working memory deficits in language performance experienced by individuals with aphasia. The significant differences in mean scores on the ASMT between the aphasia groups and the reference groups highlight the impact of aphasia on cognitive functions, particularly in tasks requiring the retention and manipulation of verbal information. These results are in agreement with available literature on aphasia, drawing attention to the importance of highlighting the main difficulties of aphasic patients in auditory memory tasks.

The overall results of this study have significant clinical implications for the proper diagnosis, cognitive assessment, therapy, and treatment of working memory impairments in individuals with aphasia. Given the observed deficits in both visual and auditory working memory, clinicians should consider introducing targeted memory training exercises in rehabilitation programmes. These interventions should be tailored to address the specific challenges faced by bilingual individuals with aphasia, taking into account the cognitive load associated with managing multiple languages. Moreover, the variability in

performance within the bilingual aphasia group underscores the importance of individualised assessment and intervention strategies. Clinicians should be aware of the potential for significant differences in cognitive processing among bilingual individuals with aphasia, and assessments should be designed to capture these individual differences accurately. Individuals with aphasia could be better supported if personalised, fit-for-purpose approaches were introduced to overcome the specific cognitive challenges.

5. Conclusion

In conclusion, the findings of this study underscore the intricate relationship between bilingualism, aphasia, and working memory. Both bilingual and monolingual individuals with aphasia demonstrated impairments in the visual and auditory working memory. These results underscore the importance of considering individual differences in cognitive processing when assessing and treating individuals with aphasia. By adopting a comprehensive approach that takes into account the unique challenges faced by bilingual individuals, clinicians can improve therapeutic outcomes and support individuals in their recovery. Further research is needed to explore the intricacies of working memory in bilingual populations affected by aphasia, ultimately contributing to a more comprehensive understanding of cognitive functions in this context.

6. Limitations

While this study provides valuable insights into working memory impairments in bilingual and monolingual individuals with aphasia, the results, along with further assumptions and suggestions, are not without limitations. One reason is the relatively small sample size, which limits the generalizability and broader applicability of the findings. Additionally, the lack of statistically significant differences between the mono-and bilingual aphasia groups reflects the need for further research with larger samples to explore these trends more comprehensively. Future studies should consider longitudinal designs to assess changes in working memory performance over time, particularly in response to targeted interventions. Furthermore, research exploring the neural correlates of working memory in bilingual individuals with aphasia may provide deeper insights into the cognitive processes involved. A proper design of behavioural tasks may be supported by a better understanding of the underlying mechanisms in individuals with aphasia, which contributes to developing more effective therapeutic approaches that cater to the unique and special needs of bilingual individuals with aphasia.

References

- Adrover-Roig, D., Galparsoro-Izagirre, N., Marcotte, K., Ferré, P., Wilson, M. A., & Inés Ansaldi, A. (2011). Impaired L1 and executive control after left basal ganglia damage in a bilingual Basque–Spanish person with aphasia. *Clinical Linguistics & Phonetics*, 25(6–7), 480–498. doi: <https://doi.org/10.3109/02699206.2011.563338>
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829–839. <https://doi.org/10.1038/nrn1201>.
- Chapey, R. (2008). Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders, Chapter 33, 879–962. Baltimore, MD: Lippincott Williams & Wilkins.
- Christensen, S. C., & Wright, H. H. (2010). Verbal and non-verbal working memory in aphasia: What three *n*-back tasks reveal. *Aphasiology*, 24(6–8), 752–762. <https://doi.org/10.1080/02687030903437690>
- Christensen, S. C., Wright, H. H., & Ratiu, I. (2018). Working memory in aphasia: Peeling the onion. *Journal of Neurolinguistics*, 48, 117–132. doi: <https://doi.org/10.1016/j.jneuroling.2018.02.001>
- Cseh, K., & Hegyi, Á. (1995). *Gyakorlatok az afázia kognitív nyelvi terápiájához*. Budapest, Hungary: Nemzeti Tankönyvkiadó.
- Dodge, Y. (2008). Kolmogorov-Smirnov Test in *The Concise Encyclopedia of Statistics*. Springer, New York, NY doi: https://doi.org/10.1007/978-0-387-32833-1_214
- Gitterman, M. R., Goral, M., & Obler, L. K. (Eds.) (2012). *Aspects of Multilingual Aphasia*. Bristol, Buffalo, Toronto: Multilingual Matters. ISBN-13: 978-1-84769-754-7
- Goetz, C. G. (2007). Textbook of Clinical Neurology (3rd Ed.). Philadelphia, PA: Saunders. ISBN 978-1-4160-3318-0.
- Kirkwood, M.W., Weiler, M. D., Bernstein, J. H., Forbes, P. W., & Waber, D. P. (2001). Sources of Poor Performance on the Rey- Osterrieth Complex Figure Test among children with learning difficulties: a dynamic assessment approach. *The Clinical Neuropsychologist*, 15(3), 345–56. doi: <https://doi.org/10.1076/clin.15.3.345.10268>
- Lee, B., Moon, H. I., Lim, S. H., Cho, H., Choi, H., & Pyun, S.-B. (2016). Recovery of language function in Korean-Japanese crossed bilingual aphasia following right basal ganglia hemorrhage. *Neurocase*, 22(3), 300–305. doi: <https://doi.org/10.1080/13554794.2016.1141966>
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). Neuropsychological Assessment (5th ed.). Oxford University Press.
- Mayer, J. F., & Murray, L. L. (2012). Measuring working memory deficits in aphasia. *Journal of Communication Disorders*, 45(5), 325–339. doi: <https://doi.org/10.1016/j.jcomdis.2012.06.002>
- Mooijman, S., Schoonen, R., Roelofs, A., & Ruiter, M. B. (2022). Executive control in bilingual aphasia: A systematic review. *Bilingualism: Language and Cognition*, 25(1), 13–28. doi: <https://doi.org/10.1017/s136672892100047x>
- Murray, L.L., Ramage, A. E., & Hopper, A. (2001). Memory impairments in adults with neurogenic communication disorders. *Seminars in Speech and Language*, 22 (2), 129–136. doi: <https://doi.org/10.1055/s-2001-13937>
- Nikraves, M., Aghajanzadeh, M., Maroufizadeh, S., Saffarian, A., & Jafari, Z. (2021). Working memory training in post-stroke aphasia: Near and far transfer effects. *Journal of communication disorders*, 89, 106077. doi: <https://doi.org/10.1016/j.jcomdis.2020.106077>
- Noether, G. E. (1963). Note on the Kolmogorov statistic in the discrete case. *Metrika*, 7(1), 115–116. doi: <https://doi.org/10.1007/bf02613966>
- Penn, C., Frankel, T., Watermeyer, J., & Russell, N. (2009). Executive function and conversational strategies in bilingual aphasia. *Aphasiology*, 24(2), 288–308. doi: <https://doi.org/10.1080/02687030902958399>
- Penn, C., Barber, N., & Fridjhon, P. (2017). Early recovery profiles of language and executive functions after left hemisphere stroke in bilingualism. *Aphasiology*, 31(7), 741–764. doi: <https://doi.org/10.1080/02687038.2016.1258538>

- Potagas, C., Kasselimis, D., & Evdokimidis, I. (2011). Short-term and working memory impairments in aphasia. *Neuropsychologia*, 49(10), 2874–2878. doi: <https://doi.org/10.1016/j.neuropsychologia.2011.06.013>
- Slakter, M. J. (1965). A Comparison of the Pearson Chi-Square and Kolmogorov Goodness-of-Fit Tests with Respect to Validity. *Journal of the American Statistical Association*, 60(311), 854–858. doi: <https://doi.org/10.1080/01621459.1965.10480833>
- Szabó E., & Sándor, K. (2021). *Afáziás feladatgyűjtemény I.* ISBN 978-615-81597-8-4 Retrieved from: https://www.efiportal.hu/wp-content/uploads/2021/07/afaziaterapias-feladatgyujtemeny-i_.pdf
- Virtanen, P., Gommers, R., Oliphant, T.E. *et al.* (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods*, 17(3), 261–272. doi: <https://doi.org/10.1038/s41592-019-0686-2>
- Walsh, J. E. (1963). Bounded probability properties of Kolmogorov-Smirnov and similar statistics for discrete data. *Annals of the Institute of Statistical Mathematics*, 15(1), 153–158. doi: <https://doi.org/10.1007/bf02865912>
- Wright, H. H., & Shisler, R.J. (2005). Working memory and aphasia: theory, measures, and clinical implications. *American Journal of Speech-Language Pathology*, 14 (2), 107-118. doi: [https://doi.org/10.1044/1058-0360\(2005/012\)](https://doi.org/10.1044/1058-0360(2005/012)) .
- Zakariás, L., Salis, C., & Wartenburger, I. (2018). Transfer effects on spoken sentence comprehension and functional communication after working memory training in stroke aphasia. *Journal of Neurolinguistics*, 48, 47-63. doi: <https://doi.org/10.1016/j.jneuroling.2017.12.002>

MERGA FEYERA WEKJIRA¹ – ANAT STAVANS²

¹University of Pannonia, Multilingualism Doctoral School
margaa30fayyoo@gmail.com

²University of Pannonia, Institute for Hungarian and Applied Linguistics
Stavansahun@gmail.com
<https://orcid.org/0000-0003-1000-5262>

Merga Feyera Wekjira – Anat Stavans: Profiles of multilingual agencies in educational contexts in Oromia, Ethiopia
Alkalmazott Nyelvtudomány, XXV. évfolyam, 2025/1. szám, 187–209.
[doi:http://dx.doi.org/10.18460/ANY.2025.1.010](http://dx.doi.org/10.18460/ANY.2025.1.010)

Profiles of multilingual agencies in educational contexts in Oromia, Ethiopia

Historically, multilingualism as a form of living and socialization has existed and has been documented since biblical times. This study explores the profiles of multilingual agencies within educational contexts, focusing on students, teachers, and parents. Utilizing a comprehensive, multi-section questionnaire, the research examines participants' self-reported language abilities (based on Clark's CANDO test), patterns of language use across various contexts and interlocutors (Stavans et al., 2009), attitudes toward each language and MPQ. Descriptive statistical analyses reveal widespread multilingualism among all participant groups, emphasizing the dominant role of primary and secondary languages in shaping educational and social experiences. The findings align with existing research on multilingualism, language hierarchy, and education's sociocultural impact. They highlight the cognitive, social, and economic advantages of multilingualism, emphasizing the need for inclusive language policies that recognize and support linguistic diversity. The study also identifies challenges faced by minority languages, underscoring the importance of equitable language education. Finally, it highlighted that educational settings significantly influence language use, shaping linguistic flexibility and multicultural personality traits.

Keywords: multilingual agencies, educational contexts, Oromia, Ethiopia

1. Introduction

Historically, multilingualism as a form of living and socialization has existed and has been documented since biblical times (Stavans & Hoffmann, 2015) with the mobility of speakers of different languages across geographic regions in pursuit of food, territory or trade, bringing peoples from different geographical and ethnic backgrounds together. This movement of groups, or "relocation" in modern terms, created numerous opportunities for cross-linguistic interaction. Indeed, human mobility has long been a driving force behind multilingualism, encompassing people from all walks of life—both educated and uneducated, wealthy and impoverished, in rural and urban settings, and across various regions. Kimber (2014: 150) stated, “The times we live in are characterized by the movement of people, ideas, goods, and services across the globe; language skills are thus necessary to confront the challenge of maintaining order and positive

relationships between people of all backgrounds and cultures.” Similarly, Stavans & Jessner (2022: 1) have argued that this human mobility required that people needed to function in more than one language to sustain a community and hence “[l]anguages [had to] become “currencies” that have different and changing “exchange rates” in different spheres of human interaction in the “communication market” at different times, in different places, and for different purposes.” As language(s) become the greatest invention of mankind and a precious property that connects people of the world together (Deutscher, 2005), it also became a catalyst agent in propagating multilingualism and multilingual societies (Blackledge & Creese, 2010), making “[t]heir multilingualism eventually ... a marker of their identity” (Stavans & Hoffmann, 2015: 11).

1.1. Factors affecting multilingualism in multilingual agents

Multilingualism could be affected either positively or negatively by many factors. Language use and purpose, language proficiency/ ability/ CANDO, language attitude, emotion, and personality (Tokuhama, 2003) are some of the major factors for both the development and stagnation of multilingualism. Apart from the distinction referring to different types of forces leading to multilingualism whether these are instrumentally or integratively driven by external (social) or internal (individual) linguistic needs and language use. While learners might have instrumental (external) or integrative (internal) reasons for learning a language, or maintaining their multilingualism, they can also have different reasons for using the language. House (2002) differentiates between “language for communication” and “language for identification” (terms taken from Hüllen, 1992). Multilingual people can choose the language and adjust the language to their needs by using their mother tongue to express their cultural identity and using another language only as an instrument to communicate and to understand each other (Dégi, 2012).

1.1.1. Language use with individuals and for different purposes

Many scholars believe that there are a number of factors which influence language choice and language use. The three dominant factors are: domain, interlocutors, and topic. The term "domain" refers to the idea that each language or dialect is associated with a specific role, setting, and/or group of people within society in which it includes work, family, and religious contexts (Spolsky, 2012; Weinreich, 1953). Fishman (1972) argues that domain is a good concept to look into how people use language both individually and in groups. With this regard, a language people use at home could be vary from the one they use at public spheres. As the same time, a language that minorities children use at home and neighbour could be different from that of the schools. These notions are apparent in the language domain theory of Fishman (2000) that language speakers in ethnic minority groups frequently link particular languages to particular domains.

The second dominant factor of language use and language choice is interlocutors. Studies depict that interlocutors also clearly have an impact on language use. With this regard, there are three patterns that Harris (2006) discovered while examining the language use of ethnic minority groups in the London village: with parents, with siblings and with grandparents. The one with parent was a hybrid language of majority and minority languages. The second used the majority language whereas the third used the minority language. However, Harris understood that the participants utilized their majority languages with grandparents and elder relatives to show respect and to proud their parents since the parents were successful in helping their children maintain their mother tongue.

A topic of discussion can also influence a language choice and use (Ritchie and Bhatia, 2012). Fishman (2000: 92) suggests that “Certain topics are somehow handled better in one language than in another in particular multilingual contexts.” This implies that a language used to talk about love may be different from the one to talk about sport or politics or business or anything else in a multilingual scenario. This could be dependent on different reasons. Namei (2008: 420) explained one of the reasons, “The use of the ethnic minority language is due to the speakers’ limited competence in the subject matters, or the lack of required vocabulary in the other language.” In addition to the above three determinant factors, age and gender differences (Harris, 2006; Wei, 1994) have also their own influence on a language choice and use.

1.1.2. Language proficiency/ ability/ CANDO

Language assessment has been a long-debated aspect in multilingualism and more specifically in the context of language education. The contested issues not only concern what is to be assessed or how but also the methodological perspective to be taken in terms of form versus function of language (Spolsky 1985). Moreover, whether assessment should focus on skill, ability, or competence and whether there is a valid and reliable measure (Shohamy, 2001; Canale, 1983), or whether there is one way to measure different types of knowledge in different contexts, at different ages, and in different stages of language development or learning (Kern, 2000; Echevarria et al., 2004). Inevitably, the context and purpose of assessment dictates the practices, ideologies and methodology employed as this is called for in measures of achievement, placement, progress that are required in different contexts such as school, business, healthcare, etc. Cloud et al. (2000) proposes a more comprehensive form-function view of language proficiency as the capacity to use language correctly and appropriately in both written and oral forms in a range of contexts; whereas Echevarria et al., (2004) has a more formalistic perspective:

Language proficiency is a measurement of how well an individual has mastered a language; proficiency is measured in terms of receptive and

expressive language skills, syntax, vocabulary, semantics, and other areas that demonstrate language abilities. (p. 224)

Since the 1980s, the language community has realized that tests must assess performance of authentic language functions, but those terms have yet to be satisfactorily defined and placed in an accepted theoretical model. Models have been proposed, but they turned out to be programmatic and heuristic rather than rigorous and testable. Canale (1983) was among the initial proponents stating a comprehensive conceptual framework that encompasses three aspects of language proficiency: linguistic, cognitive, and sociocultural. A language requires knowledge and proficiency with its linguistic components in order to be considered fluent. It also calls for prior knowledge, the ability to think critically and metacognitively, as well as the capacity to comprehend and apply cultural nuances, practices, and beliefs in a given setting. Proficiency in a language also entails the ability in the four language domains—speaking, writing, listening, and reading—appropriately for a range of contexts, audiences, and goals. Such abilities can be examined in terms of levels - basic, conversant, and advanced (Kern, 2000) or in terms of stages as proposed by Echevarria et al. (2004): pre-production, early production, speech emergence, intermediate fluency, and advanced fluency stages.

1.1.3. Language attitude

Language learners' attitudes towards the language and its speakers greatly influence the language learning process and the learning outcomes. Previous research and studies on attitudes and motivation in language learning (Csizér 2007) show that attitudes and motivation are strongly intertwined. Positive attitude towards the language and its speakers can lead to increased motivation, which in turn results in better learning experience and achievement. The relationship between attitude towards language and multilingualism is complex and paramount. Attitudes towards languages can influence individual language learning behavior, language maintenance, and ultimately the development of multilingualism in societies. Krashen (1982) argued that positive attitudes foster language learning when individuals with positive attitudes towards languages are generally more motivated to learn and use multiple languages. He emphasizes the importance of a low-anxiety environment and a positive affective filter for language learning to foster the conditions for a successful language learning experience. Also, understanding educators' attitude is essential to comprehending their decision-making in the classroom so that in the educational context, the language instructor plays a crucial role in promoting students' positive attitude towards the language and its learning which in turn promotes multilingualism (Haukås, 2016).

A positive attitude can be influenced by factors such as cultural appreciation, perceived usefulness of languages, and personal interest in diverse linguistic

experiences. Fishman (1966) highlights the link between language attitudes, cultural identity, and language maintenance. He emphasizes the role of ethnolinguistic vitality in sustaining minority languages making a tightly knit relation between cultural identity and the attitude towards the language. People often associate their language(s) with their cultural heritage and identity and consequently a positive attitudes towards one's own language(s) can promote language maintenance and preservation of cultural identity within multilingual communities.

In multilingual societies, attitudes towards different languages can affect language use patterns, language shift, and language revitalization efforts. Negative attitudes towards minority languages, for example, can contribute to language endangerment and decline. Hornberger's (2008) research in language policy and planning addressed issues related to linguistic diversity, bilingual education, and the role of language attitudes in shaping societal multilingualism. Within the specific educational institutional framework, attitudes towards languages play a role in language education and bilingual programs. Positive attitudes towards bilingualism and multilingualism in education can promote successful language learning outcomes and support the maintenance of multiple languages among students (Todor & Degi, 2016). Baker & Write (2017) advocate for inclusive language policies that value linguistic diversity and promote positive attitudes towards multilingualism in educational contexts and more recently translanguaging has sprung from the recognition of not only linguistic diversity but from the rich coexistence of this linguistic diversity housed in a single linguistic repertoire within the multilingual individual (Garcia & Wei, 2014).

1.1.4. Personality

It has been argued that personality dimensions “summarize a person's typical behavior” (Pervin & Cervone, 2010, p. 229) these have been categorized into five top indicators of the hierarchy (Pervin & Cervone, 2010). Although psychologists agree that personality is determined both by physiological and social factors (Furnham & Heaven, 1999), relatively less research has been carried out on the effect of external (social factors) and internal (individual factors) within the context of multilingual. Studies exploring the relation between the main personality indicators and multilingualism have argued that it hones the potential of one shaping the other and the direction as to whether personality shapes multilingualism or vice versa has been contested (Grosjean, 2014; Dewaele & van Oudenhoven, 2009; Dewaele & Wei, 2013; Wei & Hu, 2018). In fact, the causal pathway between multilingualism and personality traits could be multi-directional, where multilingualism is both the cause and the effect (i.e., being multilingual can push a person to develop a more multicultural personality or it is more likely to become multilingual if you are born with a certain personality profile (Dewaele & Wei, 2013; Dewaele & Botes, 2020).

Personality can play a significant role in various aspects of multilingualism, including language learning, language use, emotional experiences and overall proficiency. More specifically, personality traits impact language learning and use have been studied by Dewaele (2009) who argues that various personality traits, such as extraversion, neuroticism, openness to experience, agreeableness, and conscientiousness, can significantly influence how individuals approach language learning, interact in multilingual environments, and express emotions across languages. The role of emotional and psychological factors can shape the multilingual individuals' language experiences if different personality traits can influence how individuals manage language-related anxieties, cope with language challenges, and navigate linguistic and cultural diversity. Inevitably, the variability in individual differences in language acquisition and use by multilingual is borne out of personality traits that contribute to variability in language learning outcomes and proficiency levels. It has been argued that there is a dynamic interplay between personality and language (learning) because the relationship between personality and multilingualism is dynamic, bidirectional and multifaceted. Dewaele (2019) argues that personality traits not only influence language behaviors but can also be shaped by language experiences and cultural interactions. Jessner (2008) discusses the cognitive and affective aspects of multilingualism, including how personality traits influence language learning strategies and language use in diverse cultural settings. Others (MacIntyre, 1996; Dörnyei, 2014) investigate the role of motivation, personality, and social context in second language acquisition and multilingualism; and multicultural identity development.

In general, the multilingual profiles are very determinant in the development of multilingualism. As the world becomes increasingly interconnected, the importance of multilingual profiles has grown, shaping individuals' identities, social interactions, and professional opportunities. Understanding how people perceive and engage with multilingualism provides valuable insights into their cultural attitudes, linguistic preferences, and the broader societal factors influencing language use. The present study aims to profile each agent group as a whole in terms of their demographic background, their patterns of language use in different interlocutor contexts and for different purposes, their subjective holistic and specific evaluation of their language abilities and skills in each language, their attitude towards monolingualism and multilingualism, and their personality traits that may be related to their multilingualism since there is no prior study about it as far as the researchers knowledge was concerned. Accordingly, the study was designed to answer the following basic question: which features characterize the profile of each the multilingualism agents' group in terms of demographic (age, languages, education), proficiency in the language/s (cand-do and general), language/s use with interlocutors (intimate and formal) for different purposes (entertainment, sustainability and business),

attitude towards mono/multilingualism and personality traits? We hypothesized that: Agents with demographic features (such as age, education, and exposure to languages) who use more languages with interlocutors of the intimate and formal circles and use different languages for different life-purposes, will have a positive attitude towards multilingualism, will evaluate their language abilities in more than one language higher, and will have enhanced personality features.

2. The study

2.1. Participants

The subjects of the study were students, teachers and parents of the students. 80 students, 40 teachers and 80 parents of the students from both public and private schools of Oromia region, Ethiopia took part in the study. The researchers focused on 12th grade students who were soon to complete their basic schooling, assumed that they would soon be at the brink of moving into society either directly to the workforce or to a professional/vocational capacity building training period. The focus on the teachers, both language teachers as well as other subjects matter teachers, was to map out the perceptions and practices of these educational agents who had a pivotal role in motivating and enabling the best outcome of a multilingual individual that would be productive and constructive within the nation and outwards. Focusing on the parents provided the completion of all the active agents in the process of fostering, enabling and encouraging multilingual practices grounded in ideologies, emotions, aspirations and needs for the family, the community and the individual.

2.2. Materials

2.2.1. Questionnaire

The study employed a multiple sections questionnaire. This includes: (i) general language skills and CANDO test for self-report on language ability (a subjective evaluation of skills in each language– i.e. speaking, understanding, reading and writing) (Clark 1981), (ii) a questionnaire of language use with different interlocutors and a questionnaire of language use for different purposes (Stavans et al., 2009), (iii) a questionnaire of the attitude toward each language (Stavans et al., 2009), and (iv) a questionnaire of multilingual personality traits (MPQ) (Oudenhoven & der Zee, 2000). The researchers used these questionnaires in the study to comprehensively assess various aspects of language ability, use, attitudes, and personality traits related to multilingualism. General language skills and CANDO test were used to gauge participants' self-perceived proficiency in different languages; a questionnaire of language use with different interlocutors and a questionnaire of language use for different purposes were used to analyze patterns of language choice depending on social context and to examine how participants use different languages in various functional domains; a questionnaire of the attitude toward each language were used to measure participants' emotional,

cultural, or practical attitudes toward the languages they use; MPQ were used to explore personality traits that may be influenced by or influence multilingualism. The collected data were digitized using Google Forms for efficient organization and processing. The data were then exported to Excel and analyzed using the Statistical Package for the Social Sciences (SPSS). The analysis included descriptive statistics such as frequencies to summarize the data.

3. The results

3.1. Demographic information

The participants of this study included 80 students who were completing or just completed 12th grade in the private or public school system in Oromia, 40 in-service teachers of different subject matter working in either a private or a public school, and 80 parents of pupils schooled in the same schools in Oromia. Table 1 summarizes the participants' information in the study.

Table 1. Participants agency groups by age, gender, school type and years of education

Group	School Type		Males (N)	Females (N)	Age (years)				Years of Education		
	Private (N)	Public (N)			Average	SD	min	max	Average	min	max
Student	41	39	17	63	19.18	0.94	17	21	12	12	15
Teacher	20	20	34	6	38.53	4.87	28	51	15	15	21
Parent	40	40	61	17	47.88	5.71	36	61	10	7	17

Table 1 shows the general demographic information of the targeted agents group irrespective of school type. The agent group consisted of the same number of participants from the private and public school type in Oromia. The distribution of female and male participant was uneven as in the student agent group females were more engaging than males while in the teacher and parent agent groups male participants predominated. Unsurprisingly, students were younger than teachers and parents but unlike the expected age of graduation (typically around 18-19) the participants in this group consisted of students who were at the time of study collection starting 12th grade and those who just graduated or graduated a year before. The range of age in this group was conceptually motivated so as to have a broad perspective of the effect of schooling during and after completion of studies and on the brink of entering the workforce of the region. Parents unlike teachers were roughly 10 years older than the teachers and as the groups consisted of a male majority, this does not come as a surprise. In the present generation, the number of females has surpassed the number of males across all regions, and that is why the number of female students is greater than the males. However, when it comes to parents and teachers, due to societal traditions, males continue to receive more opportunities than females to engage in social matters. In all the years of education of teachers are higher than parents as teachers not only complete 12 years of schooling but also go to university to get an undergraduate degree and

then continue to acquire a teaching diploma/certificate. The variability in parents educational background is rather broad as some parents are highly educated (17 years of education) while others are basically educated with 10 years of schooling.

Table 2. Language repertoire by agent participant groups' Bilingualism, Trilingualism and the role of languages

Group	Reported number of languages % (N)			Afan Oromo % (N)			Amharic % (N)			Tigrinia % (N)			Guaragigna % (N)			English % (N)		
	1-2 languages	3+ languages	no answer	L1	L2	L3	L1	L2	L3	L1	L2	L3	L1	L2	L3	L1	L2	L3
Student	1.25 (80)	98.75 (80)	0 (80)	82.28 (79)	17.24 (87)	0.00 (80)	11.39 (79)	82.76 (87)	0.00 (80)	3.80 (79)	0.00 (87)	0.00 (80)	2.53 (79)	0.00 (87)	0.00 (80)	0.00 (79)	0.00 (87)	100.00 (80)
Teacher	27.5 (40)	72.5 (40)	0 (40)	63.16 (38)	6.25 (32)	0.00 (40)	23.68 (38)	93.75 (38)	0.00 (40)	10.53 (38)	0.00 (32)	0.00 (40)	2.63 (38)	0.00 (32)	0.00 (40)	0.00 (38)	0.00 (32)	100.00 (40)
Parent	25 (80)	71.25 (80)	3.75 (80)	78.75 (80)	13.16 (76)	1.25 (80)	10.00 (80)	85.53 (76)	0.00 (80)	3.75 (80)	1.32 (76)	1.25 (80)	7.50 (80)	0.00 (76)	0.00 (80)	0.00 (80)	0.00 (76)	76.25 (80)

Table 2 illustrates each agent group by whether they speak 1-2 languages (mono/bilinguals) or 3+ languages (Trilinguals). Furthermore, the report of each agent group regarding the role (L1, L2 or L3) of most frequently reported language (Afan Oromo, Amharic, Tigrigna, Guragigna and English languages). Most (98.75%) of the student agent group are multilingual speaking three or more languages, and most of them (82.28%) report Afan Oromo as their L1 while Amharic language is the L2 for 82.76% of them. Among the teacher agent group, 72.5% report they have three or more languages, where Afan Oromo is the L1 for 63.16% of them and Amharic is their L2 93.75%. The parents agency group report vastly (71.25%) to speak at least three languages, with Afan Oromo as the L1 for 78.75% of them while Amharic is a close companion in their linguistic toolkit as L2 (85.53%). A significant proportion of participants across all groups are multilingual, particularly trilingual. Afan Oromo predominantly is reported to be the L1 across all groups, with Amharic being the predominantly common L2.

Table 3a. Average holistic proficiency assessment (spoken and written skills) in each language as reported by agent participants of as a whole and by school-type

		Afan Oromo		Amharic		Tigrinia		Guaraginia		English	
		$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
Student	Private	98.17	11.71	96.34	13.18	7.93	26.48	4.27	13.58	98.17	11.71
	Public	97.44	16.01	96.79	16.40	5.13	22.35	0.00	0.00	97.44	16.01
	Both	97.81	13.89	96.84	14.64	6.65	24.58	1.58	8.34	97.78	13.98
Teacher	Private	60.00	50.26	100.00	0.00	15.00	36.63	0.00	0.00	100.00	0.00
	Public	70.00	47.02	100.00	0.00	5.00	22.36	5.00	22.36	100.00	0.00
	Both	65.00	48.30	100.00	0.00	10.00	30.38	2.50	15.81	100.00	0.00
Parent	Private	78.13	32.12	78.75	33.76	10.00	30.38	5.00	18.95	61.25	47.69
	Public	81.88	32.02	83.75	32.79	2.50	15.81	6.25	20.22	73.75	43.47
	Both	80.00	31.92	81.25	33.16	6.25	24.36	5.63	19.48	67.50	45.77

Table 3a presents the average score of the participants' holistic assessment of language ability across both agent groups and within each group by the school type for each languages. The students agent group show very high and consistent proficiency in Afan Oromo (97.81% in both types of schools, 98.17% in private and 97.44% in public schools), where Amharic and English exhibit similar scores; this is not the case for Tigrigna (6.65% in both types of schools, 7.93% in private and 5.13% in public schools) and Guragigna (1.58% in both types of schools, 4.27% in private and 0.0% in public schools). Similarly, the teachers agent group report perfect proficiency in Amharic and English languages with both scoring 100, but their proficiency in Afan Oromo is noticeably lower (65.00% in both types of schools, 60.0% in private and 70.0% in public schools). Moreover, the teachers agent group like the students indicate a low proficiency in both Tigrigna and Guragigna languages particularly for Guragigna language (2.50). The parents agent group's scores are moderate across most languages with Afan Oromo and Amharic language scoring around 80% and 81%, respectively indicating diverse proficiency levels irrespective of school type and similar scores in private school of around 78% in both languages and in public schools around 81% in Afan Oromo and 83% in Amharic. Their scores for Tigrigna and Guragigna languages are similar to those of students and teachers with Tigrigna language scoring 6.25 and Guragigna language at 5.63. Yet, their English language proficiency is lower than for teachers and students, with an average of 67.50 irrespective of their children's school type and 61.25% in the private schools as opposed to 73.75% in the public schools showing a slight variability among parents according to the school type their child is enrolled.

Table 3b. Average language-related subjective proficiency assessment (CAN-DO in speaking, understanding, reading and writing) in each language as reported by agent participants as a whole and by school-type

		Afan Oromo		Amharic		Tigrinia		Guaraginia		English	
		$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
Student	Private	98.17	11.71	96.34	13.18	7.93	26.48	4.27	13.58	98.17	11.71
	Public	97.44	16.01	96.79	16.40	5.13	22.35	0.00	0.00	97.44	16.01
	Both	97.91	11.15	98.10	7.43	8.86	28.42	4.66	0.89	95.31	11.31
Teacher	Private	60.00	50.26	100.00	0.00	15.00	36.63	0.00	0.00	100.00	0.00
	Public	70.00	47.02	100.00	0.00	5.00	22.36	5.00	22.36	100.00	0.00
	Both	70.89	42.50	100.00	0.00	10.00	30.38	5.00	0.00	100.00	0.00
Parent	Private	78.13	32.12	78.75	33.76	10.00	30.38	5.00	18.95	61.25	47.69
	Public	81.88	32.02	83.75	32.79	2.50	15.81	6.25	20.22	73.75	43.47
	Both	88.85	23.23	85.90	29.44	7.06	25.00	3.29	2.08	67.38	41.46

Table 3b provides the language related ability in specific activities (Can-Do) across groups agents, languages and school type in terms of averages and standard deviations. The findings that emerge regarding the student agents irrespective of school type is that Afan Oromo and English are reported to have the highest

performance rating, indicating these languages are dominant or preferred; whereas Tigrinya and Guragina have significantly lower performance, suggesting limited usage or proficiency among students. Moreover, the differences between private, public, and both educational settings are minimal for Afan Oromo and English, but there are variations in Tigrinya and Guragina, likely reflecting accessibility or linguistic diversity in certain school types. As for the teachers agent profile, a perfect performance rating (100%) is reported in Amharic across either or both school types suggesting it may be universally essential or mandated in the teaching environment; likewise in English suggesting its importance as a medium of instruction or prestigious language. This was not the case for the official regional language of Afan Oromo where teachers exhibit moderate performance with slight variation across all school types; and to a more extreme case, the performance rating on both Tigrinya and Guragina was very low, indicating limited need for these languages within the scholastic contexts. The parents agent groups of children attending the different school type indicated that Afan Oromo and Amharic are predominantly used, with slightly higher means in parents whose children attend public schools compared to those of private schools; however, irrespective of school type, their performance use in English is lower than that of their students and teachers cohorts suggesting a less frequent need to deploy the language as well as a lower ability in it. Moreover, Tigrinya and Guragina have very low scores overall, consistent with findings for other groups.

Table 4a. Average percent of agent participants who use each language for communication purposes with (intimate and formal) interlocutors by each school and by both combined

Agent	School Type	Afan Oromo		Amharic		Tigrigna		Guragigna		English	
		$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
Intimate											
Student	Private	87.80	31.68	87.20	31.19	4.27	16.68	6.71	24.38	35.37	39.12
	Public	84.62	36.55	41.67	48.10	5.13	22.35	0.00	0.00	0.00	0.00
	Both	86.25	33.96	65.00	46.15	4.69	19.53	3.44	17.67	18.13	33.04
Teacher	Private	60.00	50.26	52.50	49.93	10.00	30.78	5.00	22.36	12.50	31.93
	Public	63.75	48.31	45.00	51.04	5.00	22.36	5.00	22.36	0.00	0.00
	Both	61.88	48.70	48.75	49.98	7.50	26.67	5.00	22.07	6.25	23.17
Parent	Private	80.00	36.34	41.88	41.37	5.63	22.28	4.38	19.52	9.79	23.33
	Public	77.50	42.29	42.50	44.29	2.50	15.81	8.75	27.47	6.25	16.75
	Both	78.75	39.20	42.19	42.58	4.06	19.26	6.56	23.78	8.02	20.26
Formal											
Student	Private	88.78	30.35	89.27	30.03	2.44	11.13	1.95	8.72	28.78	20.52
	Public	86.67	32.87	52.31	49.97	3.08	14.17	0.51	3.20	9.23	16.45
	Both	87.75	31.42	71.25	44.76	2.75	12.63	1.25	6.63	19.25	20.97
Teacher	Private	60.00	50.26	80.00	36.71	4.00	12.31	1.00	4.47	23.00	19.76
	Public	69.00	46.56	59.00	47.01	4.00	17.89	5.00	22.36	15.00	19.33
	Both	64.50	48.04	69.50	42.96	4.00	15.16	3.00	16.04	19.00	19.72
Parent	Private	87.50	27.43	80.50	33.74	5.00	22.07	1.50	7.00	26.75	26.74
	Public	81.00	29.68	81.00	32.01	1.50	9.49	5.50	18.11	29.00	18.65
	Both	84.25	28.59	80.75	32.68	3.25	16.97	3.50	13.79	27.88	22.93

Table 4a the average percent and standard deviation of participants in each agent group (Students, Teachers, and Parents) who reported using one of the languages (Afan Oromo, Amharic, Tigrinya, Guragina, and English) as dominant for communicating with different (intimate and formal) interlocutors in different school types (Private, Public, and Both). Three aspects are reflected by these findings: the interlocutor context, the agent group and the school type impact. The intimate interlocutor context overall determines the use of Afan Oromo and Amharic as dominant across all roles, with a higher reported use among agents of private schools; whereas English is reported to be used less frequently with the intimate interlocutors and varies highly among the groups. Surprisingly, Tigrinya and Guragina have minimal usage. In the formal context, Amharic predominates across all agent groups, particularly among teachers and parents; English is relatively more dominant in formal contexts compared to intimate ones, especially among private school students and parents. Unsurprisingly, Tigrinya and Guragina remain marginally used, reflecting limited need in either context.

The particularities of language use in interlocutor context by the different agent group shows some interesting insights. Students agents, when communicating with an interlocutor of the intimate context, predominantly use Afan Oromo, especially in the private compared to the public schools, and to a more moderate extent in these contexts they report using English and to a negligible extent the marginalized languages (Tigrinya and Guragina). The trend of the use of the language by the student agents with formal interlocutors is quite similar to that of the intimate with a slightly greater use of English by all groups; yet Amharic usage increases significantly in private and combined school types. The teachers agents use to a moderate extent Afan Oromo and Amharic in the intimate context especially in the private schools whereas English is rarely used and the marginalized languages have a minimal presence. Within the formal interlocutor context, Afan Oromo and Amharic remain dominant, English usage is reported by slightly more teachers in the private school suggesting it has a role either in the instructional or professional development interactions. As for the parents agent group, in the intimate context, Afan Oromo and Amharic dominate and though English is used less frequently its use varies greatly among the parents compared to the students and teachers. Here too the marginalized languages remain absent. Within the formal context, parent agents the use of Amharic increases substantially while Afan Oromo remains the same as with the intimate context and English is used moderately but there is greater variability among the parents agents reflecting diversity in proficiency, need and exposure to the language. The third aspect that relates to the use of languages in different interlocutors contexts is influenced by the school type affiliation of the agent group. In the private schools Amharic and English are reported frequently to be used in both formal and intimate contexts suggesting an inclination to gravitate towards a more global and official languages whereas Afan Oromo is present but most dominantly in the

intimate context. In the public schools Afan Oromo is consistently reported as used by all agent group reflecting its important local presence, however, English is used less frequently compared to the private schools with a high variability of usage by the agent groups especially in the intimate context.

Table 4b. Average percent of participants who use each language for communication purposes in different contexts (in the Entertainment, Sustainability, and Business contexts) by groups

Groups	School type	Afan Oromo		Amharic		Tigrigna		Guragigna		English	
		$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
Entertainment Purposes											
Student	Private	85.85	33.24	80.49	36.74	16.10	21.55	25.37	23.25	38.05	33.41
	Public	79.49	36.27	40.00	45.42	7.18	22.71	5.13	13.55	3.72	9.30
	Both	82.75	34.68	60.75	45.72	11.75	22.43	15.50	21.58	21.31	30.08
Teacher	Private	60.00	50.26	68.00	46.07	21.00	35.23	18.00	28.21	22.00	30.37
	Public	65.00	48.94	49.00	50.46	10.00	22.94	10.00	18.92	8.00	13.61
	Both	62.50	49.03	58.50	48.65	15.50	29.87	14.00	24.05	15.00	24.28
Parent	Private	93.00	22.44	82.50	35.72	7.00	23.77	3.50	16.88	33.00	40.40
	Public	87.00	32.20	86.00	33.34	7.50	17.36	11.00	18.09	31.50	32.62
	Both	90.00	27.74	84.25	34.38	7.25	20.68	7.25	17.79	32.25	36.49
Sustainability Purposes											
Student	Private	82.44	33.23	70.24	33.80	4.39	19.75	1.95	7.49	17.07	18.74
	Public	76.92	37.99	34.87	40.45	4.62	20.24	1.03	4.47	4.10	8.18
	Both	79.75	35.51	53.00	41.01	4.50	19.87	1.50	6.18	10.75	15.89
Teacher	Private	60.00	50.26	62.00	42.50	14.00	34.40	0.00	0.00	11.00	15.18
	Public	65.00	48.94	47.00	48.68	5.00	22.36	5.00	22.36	3.00	7.33
	Both	62.50	49.03	54.50	45.74	9.50	29.00	2.50	15.81	7.00	12.44
Parent	Private	84.00	27.99	70.00	34.19	5.50	22.18	3.00	14.00	24.50	30.88
	Public	79.00	35.07	63.00	30.23	2.00	12.65	5.50	17.53	19.00	13.55
	Both	81.50	31.63	66.50	32.26	3.75	18.03	4.25	15.81	21.75	23.85
Business Purposes											
Student	Private	84.88	33.99	82.93	35.65	4.39	18.17	3.41	14.07	18.54	19.18
	Public	76.92	38.54	41.03	45.18	4.62	20.24	3.08	13.41	7.18	13.37
	Both	81.00	36.27	62.50	45.49	4.50	19.09	3.25	13.67	13.00	17.46
Teacher	Private	60.00	50.26	70.00	47.02	15.00	36.63	0.00	0.00	11.00	16.51
	Public	65.00	48.94	50.00	51.30	5.00	22.36	5.00	22.36	5.00	8.89
	Both	62.50	49.03	60.00	49.61	10.00	30.38	2.50	15.81	8.00	13.44
Parent	Private	91.00	24.79	81.00	36.22	5.00	22.07	3.00	14.00	28.50	34.16
	Public	86.00	33.03	87.00	33.45	2.00	12.65	3.50	13.50	23.50	21.19
	Both	88.50	29.13	84.00	34.77	3.50	17.94	3.25	13.67	26.00	28.36

Table 4b provided the average and SD of participant score in the use of five languages (Afan Oromo, Amharic, Tigrigna, Guragigna, and English) for three purposes (entertainment, sustainability and business) by each agent groups (Students, Teachers, and Parents), distinguishing between private, public, and both school types. The findings show that the languages used for entertainment purpose, Afan Oromo dominates among private school students and parents with high averages (85.85% and 93.00%, respectively) whereas public school

participants show slightly lower values. Amharic is consistently used by more than 40% of the cases across all agent groups and school types, and English is more significant in the private school participants (students 38.05%, parents 33.00%) but low in public schools cohorts. For sustainability purposes, the findings indicate that Afan Oromo remains prominent, especially among private school students (82.44%) and parents (84.00%); and Amharic shows a steady presence, particularly in private school settings but English usage declines significantly in public schools, especially among teachers and students (below 5%). Lastly, for business purposes the patterns are similar to entertainment, with Afan Oromo and Amharic consistently leading across all groups, moreover, English is moderately used among private school agent groups but scarcely used in public school contexts. If we profile the language use purpose by agent group, the findings show that among the student agents in the private school framework Afan Oromo and Amharic more extensively used for all purposes compared to the public school students. The teacher agent groups showed a more balanced use of language for the different purposes but the languages were less diverse with lower percentages for Tigrigna and Guragigna. The parents agent group in the private school exhibit the highest usage percentages for most languages, indicating more diverse communication habits. In all measures, the high SD values, especially in languages like Amharic and Tigrigna, reveal diverse usage patterns, likely due to regional or cultural differences.

Table 5. Average percent of participants with a favorable (Pro) and unfavorable (Con) attitude towards monolingualism and multilingualism by agent group and school type

		Monolingualism				Multilingualism				Participants in group
		Positive attitude		Negative attitude		Positive attitude		Negative attitude		
		Agree (%)	Disagree (%)	Agree (%)	Disagree (%)	Agree (%)	Disagree (%)	Agree (%)	Disagree (%)	N
Student	Private	7.32	92.68	92.68	7.32	90.24	9.76	7.32	92.68	41
	Public	12.82	87.18	100.00	0.00	100.00	0.00	12.82	87.18	39
	Both	10.00	90.00	96.25	3.75	95.00	5.00	10.00	90.00	80
Teacher	Private	20.00	80.00	85.00	15.00	85.00	15.00	15.00	85.00	20
	Public	15.00	85.00	85.00	15.00	85.00	15.00	10.00	90.00	20
	Both	17.50	82.50	85.00	15.00	85.00	15.00	12.50	87.50	40
Parent	Private	10.26	92.31	89.74	12.82	97.44	5.13	7.69	94.87	39
	Public	20.00	80.00	95.00	5.00	90.00	10.00	10.00	90.00	40
	Both	15.00	85.00	91.25	8.75	92.50	7.50	8.75	91.25	80

Table 5 shows the language attitude profiles (favorable and unfavorable) toward monolingualism and multilingualism by agents (students, teachers, parents) and school types (public and private). The findings show insights in two intertwined axes: attitude (positive and negative) towards -lingualism (mono – and multi-) and attitude by agent group. In terms of attitude towards monolingualism, there is a low percent of participants with a positive (pro) attitude to monolingualism across

all groups, for example, private school students agent report only 7.32% agreeing with monolingualism, while their public school cohorts reach 12.82%. Teachers agents are more likely to hold a positive attitude toward monolingualism compared to students or parents, especially the private school teachers at 20%. Parents agents in public schools are the most favorable towards monolingualism with 20% agreeing like the teachers in the private school and unlike all other agents. Unlike a low positive attitude towards monolingualism, the negative attitude is overwhelming as the rejection of monolingualism is nearly universal, especially among students (e.g., 92.68% in private schools). Teachers and parents also exhibit high disagreement, though slightly less emphatically. By contrast, multilingualism is overwhelmingly favored, especially among students (in private schools 90.24% and 100% in public schools), among teachers agents there is a consistent 85% favoring multilingualism across all school types. Parents agent generally show strong support for multilingualism, especially private school parents (97.44%). The mirror image that transpires shows a very low negative attitude towards multilingualism, typically below 10% and even reaching no disagreement with multilingualism as in the case of public school students report (0%).

In terms of the attitude towards mono-multi-lingualism as profiling the agent groups, we observe that the student agent group strongly reject monolingualism and overwhelmingly support multilingualism, particularly in the public schools where disagreement with multilingualism is absent altogether. By contrast, the teachers agent group displays slightly higher acceptance of monolingualism than other groups but still strongly favors multilingualism. The parents agent group shows a balanced view, with the public school parents embracing slightly more favorably monolingualism compared to the private school parents. Moreover, if we compare the agent groups by school type, we observe variability among the groups such as the private school participants generally show more nuanced attitudes (e.g., private school parents exhibit 10.26% favor monolingualism compared to 20% among public school parents); and the public school agent groups are more polarized, with strong rejection of monolingualism and near-universal acceptance of multilingualism.

Table 6. Average agent group score of multilingual and multicultural personality characteristics (MPQ) by school type

Agent Group	School Type	Cultural Empathy		Open-mindedness		Social Initiative		Emotional Stability		Flexibility	
		$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
Student	Private	89.70	9.62	89.27	10.48	89.33	10.53	89.21	10.64	89.09	10.66
	Public	91.60	9.27	91.54	9.33	91.60	9.35	91.54	9.36	91.15	9.54
	Both	90.63	9.44	90.38	9.94	90.44	9.97	90.34	10.05	90.09	10.12
Teacher	Private	81.38	11.11	83.25	14.89	72.13	7.45	62.75	11.55	79.75	14.53
	Public	82.00	13.07	87.00	14.90	71.25	15.01	56.88	24.61	87.00	14.90
	Both	81.69	11.97	85.13	14.83	71.69	11.70	59.81	19.21	83.38	14.98
Parent	Private	81.75	14.26	81.19	15.37	80.06	13.73	76.81	16.57	81.69	14.25
	Public	88.63	11.59	88.13	12.18	88.13	12.22	87.69	12.30	87.88	11.88
	Both	85.19	13.36	84.66	14.21	84.09	13.54	82.25	15.50	84.78	13.40

The findings shown in table 6 deploy the profile of each agent group in terms of the 5 multilingual and multicultural personality indicators. The student agent group in the public schools score highest across all dimensions suggesting and enriched sociocultural learning environment, with cultural empathy (91.60%) and open-mindedness (91.54) leading, whereas the private school students score slightly lower (roughly 89%), but their scores also vary greatly (SD ~10). These findings suggest that there is exposure to diverse environments in schools, fostering empathy, openness, and proactivity and possibly a broader sociocultural engagement in the public schools'. The teachers agent group exhibit significant differences compared to students with cultural empathy (81.69%) and open-mindedness (85.13%) are their strongest traits; but emotional stability is particularly low - especially in the public schools teachers (56.88%), and flexibility in the private compared to the public school teachers is lower (79.75% and 87.00% respectively). Teachers in the public schools show a lower emotional stability which may indicate stress or challenges in managing emotional demands yet they display a high open-mindedness that might be linked to their professional roles. In the public school parents agent group scores are consistently higher across all dimensions, with cultural empathy (88.63%) and open-mindedness (88.13%) being most prominent suggesting that these may be an outcome of their children's enhanced traits that result from their engagement with diverse communities. However, the private school parents agent group exhibit moderate scores (around 81%) for most traits, and there is greater variability as compared to their cohorts in the public schools. In all, the results show that the public school agent groups generally score higher across all indicators, suggesting richer exposure to multicultural settings or community involvement.

4. Discussions

This study aimed at exploring the profiles of multilingual agencies in educational contexts in Oromia, Ethiopia. The findings exhibit the following entire points: i. There is significant multilingualism across different participant groups (students, teachers, and parents) of the school environments in which many languages (i.e. Afan Oromo, Amharic, English, Tigrigna and Guragigna) played great role; ii. The profiles of the participants highlight that they are proficient (multilingual proficiency) in more than two languages where Afan Oromo and Amharic dominate; iii) It is also found that the profile of the multilingual participants could be characterized by various approaches of language choice and use; iv) Participants hold positive attitudes towards multilingualism and negative attitudes towards monolingualism; v) There is also marginalization of minority languages; and vi) Participants' group based variation in MPQ was also portrayed throughout the study. To begin with the first, the high prevalence of multilingualism among participants, particularly students and teachers, emphasizes key aspects of modern linguistic studies and educational practices. One notable aspect is the emergence of multilingualism as a standard. The findings reveal that most participants, especially within the student and teacher groups, speak three or more languages, which provides numerous cognitive and social advantages. Bialystok (2011) suggests that multilingual individuals often demonstrate enhanced cognitive flexibility, improved executive functioning, and heightened cultural awareness, all of which are beneficial in both educational and social settings. The second one is the multilingual proficiency in which the Can-Do results shed light on participants' self-perceived ability to perform specific language-related tasks. High Can-Do scores for Afan Oromo and Amharic among students, teachers, and parents indicate that these languages are not only well understood but also actively used in various practical and social contexts. Vygotsky (1978) emphasizes the importance of social interaction and practical application in language development. In contrast, lower Can-Do scores for other languages suggest that these languages are less frequently used in everyday activities, which may contribute to reduced proficiency. This limited functional use highlights a gap between language knowledge and practical application—a distinction that Cummins (2000) identifies as crucial for understanding bilingual proficiency. Cummins explains that cognitive academic language proficiency (CALP) tends to be more developed in dominant languages due to their use in formal education, while basic interpersonal communicative skills (BICS) vary depending on the role a language plays within community and family settings.

The third was the language choice and use which depicted valuable insights into how language choices are influenced by the context of communication, ranging from intimate to formal settings. The findings reveal a consistent pattern where Afan Oromo dominates across all groups—students, teachers, and parents—in both personal and formal interactions. This prevalence suggests that

Afan Oromo functions as a primary language within these communities, supporting communication in diverse contexts. The increased use of Amharic in formal settings, particularly among parents and teachers, reflects a sociolinguistic trend where a more standardized or socially prestigious language is preferred in professional and formal scenarios. Additionally, the findings highlight significant variations in language use based on purpose. Afan Oromo remains the dominant language across all contexts, especially in entertainment and business, underscoring its versatility as a widely accepted medium for both leisure and professional interactions. However, the notable use of Amharic in formal business contexts emphasizes its role as a complementary language, serving specialized or formal purposes alongside Afan Oromo. This dual-language dynamic aligns with the concept of diglossia, where two languages or dialects are used in distinct contexts within a community. Fourthly, the study's results on attitudes reveal a strong and consistent preference for multilingualism across all groups, aligning with contemporary research highlighting the cognitive, social, and cultural advantages of being multilingual. Bialystok (2011) has shown that multilingual individuals often demonstrate enhanced executive function, greater cognitive flexibility, and superior problem-solving abilities compared to monolinguals. The widespread support for multilingualism among students, teachers, and parents reflects an awareness of these benefits, which are especially relevant in an increasingly globalized world where effective cross-cultural communication and collaboration are essential. Conversely, the study also highlights a notable rejection of monolingualism, with the majority of participants expressing unfavorable attitudes toward it. This sentiment underscores a growing recognition of the limitations of monolingualism, particularly in diverse and interconnected societies. Monolingualism is often associated with reduced access to broader cultural experiences and limited cognitive adaptability (Adesope et al., 2010). The strong opposition to monolingualism across all groups suggests a societal shift in values, where reliance on a single language is increasingly viewed as a disadvantage rather than the norm.

Fifthly, the notably lower proficiency in Tigrigna and Guragigna underscores the challenges minority languages face in maintaining their relevance within educational contexts. Minority languages are often marginalized in educational policies, leading to their gradual decline as more dominant languages take precedence (Skutnabb-Kangas, 2000). The low scores in these languages, particularly among students, point to limited opportunities for exposure and usage in both formal and informal settings. This trend is particularly concerning as it mirrors the global phenomenon of language shift, where younger generations increasingly abandon their heritage languages in favor of those perceived as offering greater social and economic advantages (Romaine, 2006). The significant variability in proficiency levels, as evidenced by large standard deviations, suggests that even within communities where these languages are spoken, fluency

is unevenly distributed. This inconsistency is likely driven by disparities in exposure, support, and opportunities for consistent language use. Finally, the MPQ results reveal significant differences between public and private school participants, shedding light on the impact of social interaction and cultural context on cognitive and emotional development. Public school students scored higher in cultural empathy and open-mindedness, reflecting the influence of diverse environments that promote interaction with peers from varied cultural and linguistic backgrounds. This aligns with Vygotsky's (1978) theory that social interaction is central to cognitive development and Banks' (2013) assertion that exposure to diversity fosters inclusivity and prepares students for a globalized world. In contrast, the variability in private school scores correspond with Nieto's (2000) observation that private schools often prioritize academic excellence over cultural pluralism, potentially limiting students' engagement with diverse perspectives. Among public school teachers, the observed low emotional stability can be understood through Bronfenbrenner's (1979) ecological systems theory, which highlights the influence of systemic factors such as institutional support and classroom demands. Challenges like larger class sizes, resource limitations, and diverse student needs often contribute to stress and reduced emotional stability in public school settings (Maslach & Leiter, 2016). However, the high open-mindedness among public school teachers aligns with Gay's (2010) research, which suggests that multicultural teaching environments encourage cultural responsiveness. Ladson-Billings (1995) further supports this by emphasizing that educators in diverse contexts develop culturally relevant pedagogies, cultivating openness as a professional trait. Public school parents also scored highly in cultural empathy and open-mindedness, reflecting the intergenerational transmission of multicultural values (Grusec & Goodnow, 1994). Engaging with diverse school communities likely encourages parents to adopt and reinforce these values, creating a positive feedback loop. On the other hand, the moderate scores and variability among private school parents align with Coleman and Hoffer's (1987) findings that private schools often draw from more homogeneous cultural and socioeconomic groups, which may limit exposure to diversity. The variability within this group could be attributed to differing levels of engagement with multicultural initiatives or community interactions. Overall, the higher scores observed across all public school agent groups support Bennett's (1993) Developmental Model of Intercultural Sensitivity (DMIS), which posits that exposure to diverse cultural contexts fosters a progression from ethnocentric to ethnorelative perspectives. Public schools, with their inclusive policies and culturally diverse populations, provide ideal conditions for cultivating such traits. Hofstede's (2001) research on cultural dimensions also supports this, suggesting that environments with greater cultural variability naturally encourage open-mindedness and empathy as individuals navigate and adapt to differences.

In general, the study depicted the profiles of the multilinguals with significant multilingualism across the participants, multilingual proficiency, multilingual language use both in formal and informal contexts and for different purposes, the positive attitude of the participants towards multilingualism and negative attitude towards monolingualism, dominance language one and language two, and different attributes of MPQ. It also portrayed needs to improvements where special consideration should be given on the MPQ of the teachers and language policy that treats every language of the country in inclusive way.

5. Conclusion

The study presented a clear picture of widespread multilingualism among students, teachers, and parents, with significant roles for both the dominant language (Afan Oromo) and a secondary language (Amharic). These findings are in line with existing theories and research on multilingualism, language hierarchy, and the role of language in education. It underscores the importance of multilingualism in educational settings, highlighting its cognitive, social, and economic benefits. The study portrayed that the participants are proficient in more than two languages. They proficiently can perform in many languages in different environments for different purposes. The study also points to the need for more inclusive language policies that recognize and support the full range of languages spoken by participants. It provides a snapshot of the current state of language proficiency among different groups and educational contexts, highlighting the dominance of certain languages and the challenges faced by minorities. It also magnified the understanding of the community towards the benefits of multilingualism in which they were characterized by highly multilingual attributes and rejection of monolingualism.

Lastly, the study also underscores the impact of educational settings on multilingual and multicultural personality traits. Public schools appear to offer a more conducive environment for fostering these traits across all agent groups, likely due to greater diversity and inclusivity in their ecosystems. Addressing variability and emotional challenges within private schools and among teachers could further enhance the sociocultural learning landscape. These insights can guide policymakers, educators, and community stakeholders in designing interventions that promote multicultural competence and emotional resilience across diverse educational contexts.

References

- Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, 80(2), 207–245. <https://doi.org/10.3102/0034654310368803>
- Baker, C., & Wright, W. E. (2017). *Foundations of bilingual education and bilingualism* (6th ed.). Multilingual Matters.
- Banks, J. A. (2013). *Multicultural education: Goals and dimensions*-Center for multicultural education.
- Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In *Education for the Intercultural Experience*.
- Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 65(4), 229–235.
- Blackledge, A. & Creese, A. (2010). *Multilingualism: A Critical Perspective*. London: Continuum.
- Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Harvard University Press.
- Canale, M. (1983). *Theoretical perspectives on communicative competence*. University of Toronto Press.
- Clark, J. L. D. 1981. "Language. Chapter 9." In *A Survey of Global Understanding: Final Report*, edited by T. S. Barrows, 25–35. New Rochelle, NY: Change Magazine Press.
- Cloud, N. H., Genesee, F., & Hamayan, E. (2000). *Literacy instruction for English language learners: A teacher's guide to research-based practices*. Heinemann.
- Coleman, J. S., & Hoffer, T. (1987). *Public and private high schools: The impact of communities*. Basic Books.
- Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Multilingual Matters.
- Csizér, K. (2007). A nyelvtanulási motiváció vizsgálata Angolul és németül tanuló diákok motivációs beállítódása a nyelvválasztás tükrében [Examining language learning motivation English and German language learners' motivational attitudes towards language choice] *Új Pedagógiai Szemle* 57(6): 54–68.
- Dégi, Z. (2012). Types of multilingualism explored in the Transylvanian school context. *Jezikoslovlje*, 13(2), 645–666.
- Deutscher, G. (2005). *The unfolding of language: An evolutionary tour of mankind's greatest invention*. Macmillan.
- Dewaele, J. M. (2009). Individual differences in second language acquisition. *The new handbook of second language acquisition*, 2, 623–646.
- Dewaele, J. M. (2019). The effect of classroom emotions, attitudes toward English, and teacher behavior on willingness to communicate among English foreign language learners. *Journal of Language and Social Psychology*, 38(4), 523–535.
- Dewaele, J. M., & Botes, E. (2020). Does multilingualism shape personality? An exploratory investigation. *International Journal of Bilingualism*, 24(4), 811–823.
- Dewaele, J. M., & Van Oudenhoven, J. P. (2009). The effect of multilingualism/multiculturalism on personality: No gain without pain for third culture kids? *International Journal of Multilingualism*, 6(4), 443–459.
- Dewaele, J. M., & Wei, L. (2013). Is multilingualism linked to a higher tolerance of ambiguity? *Bilingualism: Language and Cognition*, 16(1), 231–240.
- Dörnyei, Z. (2014). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge.
- Echevarria, J., Vogt, M. E., & Short, D. J. (2004). *Making content comprehensible for English learners: The SIOP model* (1st ed.). Pearson Education.
- Fishman, J. A. (1966). *Language Loyalty in the United States: The Maintenance and Perpetuation of Non-English Mother Tongues by American Ethnic and Religious Groups*.
- Fishman, J. A. (1972). *The Sociology of Language: An Interdisciplinary Social Science Approach to the Study of Language in Society*. Rowley: Newbury House.
- Fishman, J. A. (2000). Language planning for the "other Jewish languages" in Israel: An agenda for the beginning of the 21st century. *Language Problems and Language Planning*, 24(3), 215–231.
- Furnham, A., & Heaven, P. (1999). *Personality and social behaviour*. Arnold.

- García O., Li W. (2014). *Translanguaging: Language, Bilingualism and Education*. London: García O, Li W (2014) *Translanguaging: Language, Bilingualism and Education*. London: Palgrave Macmillan Pivot. Palgrave Macmillan Pivot.
- Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice*. Teachers College Press.
- Grosjean, F. (2014). *Bilingual: Life and reality*. Harvard University Press.
- Grusec, J. E., & Goodnow, J. J. (1994). Impact of parental discipline methods on the child's internalization of values: A reconceptualization of current points of view. *Developmental Psychology*, 30(1), 4–19.
- Harris, R. (2006). *New ethnicities and language use*. Springer.
- Haukås, Å. (2016). Teachers' beliefs about multilingualism and a multilingual pedagogical approach. *International Journal of Multilingualism*, 13(1), 1–18.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Sage.
- Hornberger, N. H. (Ed.). (2008). *Encyclopedia of language and education* (Vol. 1). New York: Springer.
- House, J. (2002). Developing pragmatic competence in English as a lingua franca.
- Knapp, K., Meierkord C., eds. *Lingua Franca Communication*. Frankfurt am Main: Peter Lang, 245–267.
- Hüllen, W. (1992). Identifikationssprachen und Kommunikationssprachen. *Zeitschrift für Germanistische Linguistik*.
- Jessner, U. (2008). Teaching third languages: Findings, trends and challenges. *Language teaching*, 41(1), 15–56.
- Kern, R. (2000). *Literacy and language teaching*. Oxford University Press.
- Kimber, L. (2014). Attitudes and beliefs of students toward bi-/multilingualism at an international university in Japan. *Ritsumeikan Journal of Asia Pacific Studies*, 33(33), 139–152.
- Krashen, S. D. (1982) *Principles and Practices of Second Language Acquisition*. Oxford: Pergamon Press.
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, 32(3), 465–491.
- MacIntyre, P. D. (1996). *Affective variables in second language acquisition*. Cambridge University Press.
- Maslach, C., & Leiter, M. P. (2016). *Burnout: A multidimensional perspective*. Routledge.
- Namei, S. (2008). Language choice among Iranians in Sweden. *Journal of Multilingual and Multicultural Development*, 29(5), 419–437.
- Nieto, S. (2000). *Affirming diversity: The sociopolitical context of multicultural education*. Longman.
- Van Der Zee, K. I., & Van Oudenhoven, J. P. (2000). The Multicultural Personality Questionnaire: A multidimensional instrument of multicultural effectiveness. *European journal of personality*, 14(4), 291–309.
- Pervin, L. A., & Cervone, D. (2010). *Personality: Theory and research* (11th ed.). John Wiley & Sons.
- Ritchie, W. C., & Bhatia, T. K. (2012). Social and psychological factors in language mixing. *The handbook of bilingualism and multilingualism*, 375–390.
- Romaine, S. (2006). *Planning for the Survival of Linguistic Diversity*. Language Policy.
- Shohamy, E. (2001). Democratic assessment as an alternative. *Language testing*, 18(4), 373–391.
- Skutnabb-Kangas, T. (2000). *Linguistic Genocide in Education or Worldwide Diversity and Human Rights?* Routledge.
- Spolsky, B. (1985). The limits of authenticity in language testing. *Language Testing*, 2(1), 31–40.
- Spolsky, B. 2012. "Family Language Policy – the Critical Domain." *Journal of Multilingual and Multicultural Development* 33 (1): 3–11. doi:10.1080/01434632.2011.638072.
- Stavans, A., E. Olshtain, and G. Goldzweig. 2009. "Parental Perceptions of Children's Literacy and Bilingualism: The Case of Ethiopian Immigrants in Israel." *Journal of Multilingual and Multicultural Development* 30 (2): 111–126.
- Stavans, A., & Hoffmann, C. (2015). *Multilingualism*. Cambridge University Press.
- Stavans, A., & Jessner, U. (Eds.). (2022). *The Cambridge handbook of childhood multilingualism*. Cambridge University Press.

- Tódor, E. M., & Dégi, Z. (2016). Language attitudes, language learning experiences and individual strategies what does school offer and what does it lack? *Acta Universitatis Sapientiae, Philologica*, 8(2), 123–137.
- Tokuhamma-Espinosa, T. (Ed.). (2003). *The multilingual mind: Issues discussed by, for, and about people living with many languages*. Portsmouth, NH: Heinemann.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Weinreich U. (1953). *Languages in Contact; Findings and Problems*, New York, Publications of the Linguistic Circle of New York 1.
- Wei, R., & Hu, Y. (2018). Exploring the relationship between multilingualism and tolerance of ambiguity: A survey study from an EFL context. *Bilingualism: Language and Cognition*, 22(5), 1209–1219.

XIX. Alkalmazott Nyelvészeti Doktoranduszkonferencia

(Budapest, 2025. február 7.)

Az MTA Alkalmazott Nyelvészeti Munkabizottsága és a HUN-REN Nyelvtudományi Kutatóközpont 2025-ben is meghirdette doktoranduszoknak és mesterszakos hallgatóknak szóló, magyar és angol nyelvű alkalmazott nyelvészeti konferenciáját. A XIX. Alkalmazott Nyelvészeti Doktoranduszkonferenciát (19th Conference on Applied Linguistics for PhD Students) 2025. február 7-én rendezték meg online formában.

Az évente megrendezésre kerülő Alkalmazott Nyelvészeti Doktoranduszkonferenciák sorozat célja segíteni és támogatni az alkalmazott nyelvészeti kutatásokat, összekötni egymással a tudományterület fiatal kutatóit, valamint támogatni az alkalmazott nyelvészeti doktori képzést, és elősegíteni a hallgatók magyar és a nemzetközi tudományos életbe való bekapcsolódását. E célokkal összhangban a konferenciasorozat lehetőséget kíván teremteni az alkalmazott nyelvészet területén végzett magas szintű doktori kutatások bemutatására és a különböző témákban felmerülő kutatási kérdések és problémák megvitatására.

A 2025. évi konferenciát a HUN-REN Nyelvtudományi Kutatóközpont képviselőjében Lipp Veronika főigazgató-helyettes nyitotta meg. A megnyitó után négy szekcióban összesen tizenhét előadás hangzott el: hat magyar, tizenegy pedig angol nyelven.

A konferencia nyitószekciója egy négy előadásnak teret adó angol nyelvű szekció volt, Grácsi Tekla Etelka elnöklésével. Gyimesi Alexandra (*How to test formulaic sequences in Hungarian learners of German?*) azt vizsgálta, hogy bizonyos meghatározott tényezők, valamint a magyar és az angol nyelv hatása hogyan befolyásolja a németül tanuló magyar középiskolás diákok német nyelvi képességeit. Ezt követően Awder Raza Aziz (*Examining Iraqi Kurdish EFL learners' perceptions towards intercultural communicative competence*) egy olyan kutatást mutatott be, amely iraki kurd EFL-tanulók anyanyelvhez és angol nyelvhez kapcsolódó attitűdjét vizsgálta olyan környezetben, ahol az angol lingua francaként funkcionál. Han Htoo Naung (*EFL learners' perceptions of their teachers' explicit motivational strategies: A questionnaire study in the Hungarian higher education context*) egy kérdőíves felmérés eredményeit mutatta be, amely angolul tanuló egyetemi hallgatók attitűdjét vizsgálta a tanáraik által alkalmazott explicit motivációs stratégiák szempontjából. Mai Nelly Kaw (*Non-English majors' willingness to communicate in English after COVID-19: An interview study in Hungarian tertiary education*) pedig egy interjú vizsgálat eredményeit ismertette, amely nem angol szakos hallgatók angol kommunikációra való hajlandóságának befolyásoló tényezőit vizsgálta.

A második szekcióban, melynek elnöke Bodó Csanád volt, három angol és két magyar nyelvű előadás hangzott el. Kostic Adrianna (*Minorities in the digital*

space: Translanguaging practices of Vojvodina Hungarians online) egy a vajdasági magyarok digitális szokásait és online nyelvi gyakorlatait vizsgáló pilot kutatás eredményeit mutatta be. Nour Ben Braham (*Wesh as a mirative marker in contemporary French: A cognitive pragmatics approach*) egy arra irányuló kutatást ismertetett, hogy a francia *wesh* szó működhet-e meglepetést, váratlanságot kifejező szóként, és amennyiben igen, hogyan. Baha' Al-Deen Alnawas (*Individual differences in writing anxiety during offline and online exams: An interview study with university students in Hungary*) angol szakos egyetemi hallgatók online és offline vizsgatapasztalatai alapján vizsgálta az angol nyelv használatával kapcsolatos szorongást, az idegennyelvi-szorongás, a digitális szorongás és a vizsgahelyzetben való szorongás szempontjából. Borbély Beáta („*Vannak olyan kórházi dolgozók is, akik emberi hangnemben nyilvánulnak meg.*” – *A szülészeti ellátásban tapasztalt verbális abúzus és a pozitív kommunikációs gyakorlatok példái az anyai narratívákban*) édesanyák szülészeti ellátásban megélt kommunikációs tapasztalatait elemezte a verbális abúzusra és a pozitív kommunikációra vonatkozó valós példákon keresztül. A szekciót Kaiser Kamilla zárta (*A Találj ki új szót projekt vizsgálata az udmurt nyelvben*), aki egy udmurt nyelvtervezési projekt eredményességét vizsgáló kutatást mutatott be.

A konferencia harmadik szekciójában Markó Alexandra elnöklésével négy magyar nyelvű előadás hangzott el. Szánthó Zsuzsa (*Egy tréning hatása az idegen nyelvi szorongásra, valamint a magánhangzószerű hezitálások és a hezitációs szerepben használt diskurzusjelölők gyakoriságára*) egy idegen nyelvi szorongással kapcsolatos kutatást mutatott be, amely a tréningmódszer hatásait vizsgálta különböző hezitációs jelenségekre, spanyol nyelven legalább B2 szinten beszélő 19-23 éves nők részvételével. Hagymási Judit egy olyan kutatásról számolt be (*Csak-ot tartalmazó kifejezések értelmezésének vizsgálata óvodásoknál*), amely azt vizsgálta, mennyire értelmezik pontosan a 3-6 éves korosztályba tartozó gyermekek a *csaknem*, *igencsak* és *csak nem* kifejezések jelentését. A szekció harmadik előadója, Virág Nándor, egy arra irányuló kísérletről számolt be (*A magyar WORDNET alkalmazása főnevek szemantikai kategóriáinak definiálására*), hogy létrehozható-e nyelvtechnológiai módszerekkel egy olyan teljesen automatikus rendszer, amely főnévlistákhoz megfelelő szemantikai kategóriacímként tud rendelni. Szakács Eszter (*A szemmozgás és a munkamemória szerepe a következtetési szövegértésben idegen nyelven: Egy szemkamerás pilot vizsgálat*) pedig egy szemkamerás vizsgálat eredményeit mutatta be, amely a szemmozgás és a munkamemória szerepét vizsgálta a következtetési szövegértésben egyetemi hallgatók részvételével.

Végezetül a konferencia angol nyelvű zárószekciójában Fóris Ágota elnökletével szintén négy előadás hangzott el. Xuan Zhao (*A corpus study of figurative sense and euphemism in the case of perception verbs “taste” in Mandarin Chinese*) három ízeletet kifejező kínai ige és a hozzájuk kapcsolódó tárgy kollokációját vizsgálta azzal a céllal, hogy feltárja az ige közötti

jelentéskülönbségeket. Ezt követően Liang Yibo (*A contrastive study of cancer metaphors in English and Chinese*) a rák (mint betegség) kifejezésére használt angol és kínai eufemizmusokat elemzett a metaforák kulturális és kognitív felfogása szempontjából. Silhavy Zita (*Diachronic Analysis of Positive Face-Threatening Acts in Disney Movies: an Analysis of Princesses and “Saying No”*) pragmatikai szempontból vizsgálta, hogy mi jellemzi a női szereplők „nemet mondás” aktusának diakronikus fejlődését a Disney-filmekben, és milyen hatással lehet ez a fiatal nézők nőkről alkotott képére. A konferencia utolsó előadói, Nilay Nur Tasdemir és Ahmad Naji, pedig 180 egyetemi hallgató részvételével vizsgálták, hogy milyen asszociációkat mozgósítanak magyar, jordániai és török egyetemi hallgatók az „italok”, „érzelme”, „barátság” és „siker” szavak hallatán, és milyen interkulturális különbségek figyelhetők meg az asszociációk szempontjából (*Insights into university students’ categorization of drinks, emotions, friendship, and success in Hungary, Jordan, and Türkiye*).

Az alkalmazott nyelvészet számos különböző területéhez kapcsolódó, rendkívül színvonalas előadásokat felvonultató XIX. Alkalmazott Nyelvészeti Doktoranduszkonferenciát a munkabizottság és a konferencia programbizottságának elnöke, Fóris Ágota zárta gondolataival.

A konferencia előadói – a korábbi évek gyakorlatához hasonlóan – 2025-ben is lehetőséget kaptak előadásaik írott változatának publikálására. Az Alkalmazott Nyelvészeti Doktoranduszkonferenciák tanulmánykötete 2022-től az *Alkalmazott Nyelvtudomány* folyóirat különszámaként jelenik meg, így a 2025-ben elhangzott előadások anyagai is ebben a formában fognak megjelenni, várhatóan 2025. vége felé.

JAKAB DOROTTYA
 ELTE Eötvös Loránd Tudományegyetem
 HUN-REN Nyelvtudományi Kutatóközpont
 jakab.dorottya@nytud.hun-ren.hu
<https://orcid.org/0000-0002-2829-3103>

Ara Rauch (szerk.):
Szókereső írószótár. Magyar főnevek szókapcsolatai
 (Budapest: Tinta Könyvkiadó. 2024. 826 oldal)

A szótárak választékának palettája rendkívül széles körű: léteznek egynyelvű és kétnyelvű, vagy akár kettőnél is többnyelvű változatok. Az egy- és kétnyelvűek körébe tartoznak a szójelentést tételező értelmező szótárak, kiadásaik legkülönbözőbb terjedelmében. Ezek elsősorban az anyanyelvű nyelvhasználók számára készülnek. Nagy a választéka a nyelvet tanulóként elsajátító alanyok számára kiadott szótáraknak is. Az ún. „tanulói szótárak” zöme is szójelentés-értelmező funkciójú. Ezek mellett tucatnyi más, speciális típusú szótár létezik: szinonimaszótárak, kollokációs szótárak, kifejezés- és szólásszótárak, helyesírási útmutató szótárak, a szavak történetét és eredetét reprezentáló szótárak, terminológiai, adott szakmák profilját reprezentáló szótárak és mások.

A jelen recenzió tárgyát az Ara Rauch főszerkesztő és Kiss Gábor szerkesztő, valamint akadémiai munkacsoportjuk által a magyar lexikográfiai szakirodalomban elsőként szerkesztett magyar főnévi kollokációs szótár képezi. Az angol nyelvész, John Rupert Firth definíciója nyomán *kollokáció*-nak az egymás társaságában magas arányban előforduló lexikális egységek kapcsolatát nevezzük (Firth 1957: 11). Ennélfogva a kollokáció elsősorban morfoszintaktikai alapú, lexiko-grammatikai jelenségnek, frazeológiai státusú lexikai egységnek, szószerkezeti kapcsolatnak tartható. A korpusznyelvészet diszciplinájának megalapozója, John Sinclair nézete szerint a kollokációk a nyelvhasználók lexikális memóriájában előrekonstruált jelentéses kapcsolatok formájában rögzítettek, melyek alapját képezik az ún. idiomatikus konstrukciók nyelvhasználatának (Sinclair, 1991: 110, 173). A kollokációval foglalkozó nemzetközi szakirodalmat a lexikális kapcsolat jelentésalapú természetének kutatása (pl. az ún. szemantikai prozódia megvalósulása) mellett a diszkurzív/textuális alapú mennyiségi arányszám mutatóinak vizsgálata uralja.

Ahogy az a könyv címében is kifejezést kap, a szótár célközönségének a szerkesztők olyan, a magyart anyanyelvként beszélő egyének sokaságát tartják, akik nyelvhasználatuk beszédtevékenységük mellett kiterjed lexikális kifejezésük választékosságának igényére. Emellett ajánlják a szótárt a magyart idegen nyelvként elsajátító tanulóközönségnek is, azzal a céllal, hogy tájékozódhassanak a lexikális kifejezések használatát illetően abban, „hogyan egy adott szóhoz melyik szót kell és lehet hozzáilleszteni, hogy szépen csengő, helyes magyar szókapcsolatot hozzanak létre”. Recenzorként azonban a használati célkitűzés ilyenén megfogalmazásával, annak erőteljesen preskriptív tónusa miatt nem érthetők egyet. Ugyanis ennek megvalósíttatása helytelen módon befolyásolná, korlátozná a lexikális nyelvi reprezentáció szabadságát és nyelvhasználati kreativitását.

A „Szókereső írószótár” adatainak forrása „A magyar nyelv értelmező szótára”, melynek megjelenésére 1958 és 1961 között került sor. A hétkötetnyi szótár 58.000 címszót tartalmaz. Indokolatlan számomra, hogy Ara Rauch és munkatársai miért egy ilyen, a múlt évszázad közepén készített akadémiai szótárt választottak ki arra a célra, hogy adataiból a 21. században kollokációs szótárt szerkesszenek, mégpedig azzal a céllal, hogy általa segédanyagot biztosítsanak korunk nyelvhasználóinak, ugyanis annak szóanyaga számos, leginkább kollokációs tekintetben nem tükrözi, nem is tükrözheti a mai magyar nyelvhasználatra jellemző lexikális szókapcsolati formák állapotát. A Szókereső írószótár 6545 – a szerkesztők által az Előszóban említett (5 és 9. oldal), meglátásuk szerint a magyar nyelvben leggyakrabban megjelenő főnévi lexikális kapcsolatokkal előforduló – címszó 124.500 kollokációját tartalmazza. Ennek listáiból azonban számos, a mai nyelvhasználatban nagy gyakorisággal előforduló példa kimaradt, köztük például a magyarra oly jellemző, különféle típusú szóösszetételek. Példaként említem ezekre a hiányokra a *tűz* címszó alatt olyan összetételek kimaradását, mint a jelzős szerkezetű *tűzpiros*, a nominálisan összekapcsolt *tűzfészek*, *tűzgolyó*, *tűzkerék*, *tűzmadár*, *tűzvész*, az építészetben előforduló *tűzfal*, mely a mai számítógépes nyelvhasználat alapszókincséhez is tartozó lexikális egység. A *vér* címszó alatt nem szerepel a *vérvörös* és a *vérkép* a számos feltüntetett nominális alapú szóösszetétel mellett. Több tucatnyi hasonló kollokatív szókapcsolat reprezentációja hiányzik a kötetből. Érdekes kapcsolatot képez a *tűz* és a *vér* jelzős összetételű kollokációja a *tüzesvérű* példában, mely konstrukció, talán nem véletlenül, nem szerepel a mintegy 60 évvel ezelőtt szerkesztett forrásban. Jelzős főnévi kollokációi között kiemelt gyakorisággal rendelkezik a *tüzesvérű paripa/menyecske* konstrukció.

A szótár nem ad számot címszavainak tételezésében az igekötő által az ige alapjelentésének bővülését, tágabbá válását kiváltó mechanizmusok lexikális megnyilvánulásáról, a deverbális képzésű főnevek heterogén természetű reprezentációjáról, tudáskeret-vonzatukról (‘frame-relatedness’) és kollokációs sajátosságairól. Ennek tudományos igényű felmérésére egy, a magyar nyelvre vonatkozó kollokációs szótár lehetőséget biztosíthatna. Ennek hiánya dacára mégis a szerkesztők helyes döntése konstatalható amellett, hogy gyűjtésükben az ilyen természetű főnévi konstrukciókat külön szótári tételekként szerepeltetik (lásd többek között a *szív*, *fut*, *olvas*, *jelent* igék igekötővel ellátott deverbális nomenjeinek tételeit és azok kollokációit).

További hiányt képez a szótárban, hogy a szerkesztők a nominális alapú idiomatikus és metaforikus tartalmakat teljes mértékben figyelmen kívül hagyták. A szerkesztők részéről a szótárról általános és szakmai orientáltságú tájékoztatást nyújtó bevezető fejezetben ennek a döntésnek semmiféle indoklása nem olvasható.

A szócikkek felépítéséről a mindössze 18 oldalas Előszó ad ismertetést (8–21). Minden szócikket címszó vezet be, amit a szótár kollokációs profiljából fakadóan

természetesen nem követ a szótétel jelentéseinek hagyományos módon történő értelmező azonosítása. A címszó után maximum 6 sorszámra tételezeten szerepelnek az adott lexikális egység aspektuscsoportjai. Ezek, a szerkesztők megfogalmazása szerint „egy-egy szó nyelvbéli helyzetét” mutatják meg. Hangsúlyozzák továbbá, hogy „az egyes aspektusok között nincs éles határvonal, sok esetben a szótár szerkesztői egyéni nyelvi kompetenciájának, nyelvérzékének függvénye volt egy-egy szótári nyelvi egység, kollokáció besorolása valamely aspektuscsoportba” (15). Az aspektuscsoport fogalmának ilyen, szakmai szempontból definiálatlan megfogalmazása tulajdonképpen értelmezhetetlen a szakmai olvasó, de egyúttal az általános nyelvhasználó számára is. A kollokációk tételezésének illetően megközelítése súlyos hiba. Mind az *aspektuscsoport*, mind pedig a *nyelvbéli helyzet* fogalma definiálásra szorul. A fenti idézet, mely szerint a szótételek mögötti jelentéses csoportok megadása a szerkesztők intuitív, egyéni megítélésének függvénye, megkérdőjelezi a címszó jelentése mögötti kollokatív szókapcsolatok „aspektusbeli” viszonyának megítélését és reprezentációját a szótárban.

A szerkesztők helyesen mutatnak rá arra, hogy az általuk főképp morfoszintaktikai eszközökkel deklarált hatféle aspektuscsoport nem mindegyike reprezentált egységesen az egyes szótételek jelentésében. A kollokatív illeszkedésnek ezt a fontos tényezőjét módszeresen követték a szóegységek jelentésviszonyainak tételezésében.

További hibának ítélem meg a szerkesztők által az Előszó 21. oldalán deklarált szempontot, mely szerint a szótár „nem közli teljeskörűen a címszó enciklopédikus, ismerettári háttérét”. Ugyan a kollokációs szótárak nem jelentésmeghatározás alapján szerkesztettek, a lexikális vonzatkörök tételes reprezentációja mégis lehetőséget ad(na) arra, hogy az egyes szótételek prototípusos jelentései az enciklopédikus alapú tudáskereti (frame-based) és háttér- (scenic knowledge-based) ismeretek figyelembevételével tételeződjenek, sorrendileg konceptuálisan rögzített ismérvek alapján (lásd Andor, 1985: 212–218; 2010). A szójelentések ilyen, kollokációkra alapozott megkülönböztetésének rendkívül fontos szerepe van/lehet a mögöttes kontextuális, lexikális-pragmatikai természetű ismérvek szótári reprezentálásában. Márpedig ennek a szempontnak az érvényesítése, megjelenítése nem kapott helyet a Szókereső írószótárban. A kollokációk sorrendjét a mentális lexikonban tárolt, benne aktivált és onnan kiemelt tudáskeret-alapú tartalom kognitív szálienciájának mentén kellene tételezni. Ehhez azonban ismerni kellene ennek a lexikális faktornak a címszó jelentése szempontjából releváns asszociatív alapon megnyilvánuló potenciális tényezőit.

Alapvető problémáját képezi egy kollokációs szótárnak az a tény, hogy gyűjtött szóanyaga nem korpusznyelvészeti eljárásokra alapozott. Sajnos a Szókereső írószótár ilyen szakkönyv. Amint az a fentiekben említésre került, lexikális forrásanyaga egy mintegy 60 évvel korábban készült nyelvészeti forrásra épül,

amikor a korpusznyelvészeti diszciplínájának gyakorlatilag az alapjait sem fektették még le a kutatók. A kollokációk rendszerszerű tanulmányozására csak jóval később került sor, kezdetben az alkalmazott, majd jóval később az elméleti nyelvtudományban, a nagy volumenű nyelvi, elsősorban lexikális alapú korpuszok fejlesztése során. Ennélfogva a Bárczi Géza és Országh László szerkesztői vezetése mellett létrehozott magyar értelmező szótár teljes anyaga nem lehetett korpuszalapú, vagy akár számítógépes eszközökkel gyűjtött munka. A kollokációk vizsgálata, azaz gyűjtése és adatolása során viszont erre feltétlenül szükség lenne. A szókapcsolati előfordulások kollokációs alapú feldolgozása során elengedhetetlenül szükséges a gyűjtött szóanyag gyakorisági mutatóinak komputációs értékelése (Kosem, 2016: 82–83). Ilyen mutatók a jelen szótár lexikális tételeinek esetében sehol sem szerepelnek. Ez pedig határozott gyengesége a jelen recenzió tárgyát képező, magát kollokációs szótárnak tartó műnek.

A fentiek alapján kimondható, hogy a Szókereső írószótár nem nyerte el recenzorának tetszését. Nem tartom hasznos segédeszköznek, sem az írástevékenységet folytató, a nyelvi kifejezőmód választékosságára törekvő magyar anyanyelvű nyelvhasználó, sem pedig a nyelvtanuló számára. Ennélfogva az Előszóban leírt célközönsége nyelvi igényeit véleményem szerint nem képes teljesíteni.

Hivatkozott irodalom

- Andor József (1985). On the psychological relevance of frames. *Quaderni di Semantica*, VI/2, 212–221.
 Andor József (2010). Discussing frame semantics. *Review of Cognitive Linguistics*, 8/1, 157–176.
 Bárczi Géza & Országh László (1958–1961). *A magyar nyelv értelmező szótára*. I–VII. Budapest: Akadémiai Kiadó.
 Firth, John Rupert (1957). *Papers in Linguistics*. Oxford.
 Kosem, Iztok (2016). Interrogating a corpus. In Durkin, Philip (szerk.), *The Oxford Handbook of Lexicography* (76–93). Oxford: Oxford University Press..
 Sinclair, John (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

ANDOR JÓZSEF
 Pécsi Tudományegyetem
 Angol Nyelvészeti Tanszék
 andor.jozsef@t-online.hu

Svindt Veronika – Bóna Judit (szerk.):
Általános nyelvészeti tanulmányok XXXVI.
Tanulmányok az anyanyelv-elsajátítás köréből
(Budapest: Akadémiai Kiadó. 2024. 340 p.)

Az *Általános nyelvészeti tanulmányok* sorozat legújabb, 36. kötete Svindt Veronika és Bóna Judit szerkesztésében jelent meg. Az átfogó és összefoglaló jellegű gyűjtemény a nyelvelsajátítás témakörét alapvetően nyelvészeti szempontból közelíti meg, egybegyűjti a tudományterületen tevékenykedő hazai szakértők munkáit, bepillantást enged a nyelvelsajátítás-kutatás aktuális szakirodalmába, és bevezeti az olvasót ezen interdiszciplináris tudományterület tematikailag és módszertanilag egyaránt gazdag világába. A teljes terjedelmében a nyelvelsajátítás kérdéskörének szentelt kötet jelentős mérföldkő a sorozat történetében. A kötetbe foglalt tanulmányok a tudományterületen újonnan felmerülő kutatási kérdéseket és elméleti kérdésfelvetéseket tárgyalják.

A tizenhárom tanulmány a nyelvfejlődés különböző területeit mutatja be. A szerzők a nyelvelsajátítás különböző aspektusait, a különböző nyelvi szinteken megvalósuló folyamatokat, illetve az ezekhez kapcsolódó széles életkori spektrumon mért legfrissebb kutatási eredményeket ismertetik. Sajátjaik mellett bemutatják a témához kapcsolódó külföldi, valamint magyar vonatkozású kutatási eredményeket, ugyanakkor kijelölik a jövőbeni kutatások lehetséges és szükséges irányait is.

Babarczy Anna és Csányi Petra a nyelvfejlődés pragmatikai szintjével foglalkozik. A szerzők a kognitív rugalmasság fejlettségének, illetve a kommunikációval szembeni társas-kulturális elvárásoknak a hároméves és négyéves gyermekek metaforaértési képességének fejlettségére gyakorolt hatását vizsgálják. A vizsgálati eredmények alapján arra következtettek, hogy minimalizált kontextuális kognitív terhelés esetén az óvodás korcsoportban megnyilvánuló szó szerinti jelentéshez való ragaszkodás nem is annyira a kognitív rugalmasság hiányának, mint inkább a társas-kulturális elvárásoknak tudható be.

Csépe Valéria a prozódia fejlődésével foglalkozik, a hangsúlyészelelés és -feldolgozás idegrendszeri hátterét és kognitív feltételrendszerét, a folyamat csecsemőkori fejlődését mutatja be. Mivel a nyelvfejlődés ezen területén zajló folyamatokat az újszülöttekre jellemző korai prozódiai érzékenység nagymértékben meghatározza, a szóhangsúly feldolgozásának korai fejlődését a szerző az anyanyelvi és idegen nyelvi beszédhangsúlyra való érzékenység korai kialakulásának empirikus szakirodalmára alapján ismerteti, és ezt a folyamatot magzat-, illetve újszülöttkortól egyéves korig nyomon követi.

Dobó Dorottya, Lukács Ágnes és Lukics Krisztina Sára a statisztikai tanulás és a nyelvi fejlődés összefüggéseit vizsgálják a tipikus és az atipikus nyelvelsajátítás folyamatában egyaránt. Mivel a statisztikai tanulás és a nyelvi képességek közötti kapcsolatban az általános kognitív képességek jelentős közvetítőnek bizonyultak,

vizsgálatainkban a szerzők az ezeket befolyásoló kognitív tényezőket is figyelembe vették. Különös figyelmet fordítottak a munkamemória és a végrehajtó funkciók működésére, amelyek az anyanyelv elsajátítása szempontjából nélkülözhetetlenek.

Gósy Mária a beszédfeldolgozás mechanizmusának kialakulását és fejlődését mutatja be. Megismerhetjük a hallási, a beszédészlelési és beszédmegértési, illetve értelmezési működésmechanizmusok fejlődését és jellegzetes folyamatműködéseit, valamint a mentális lexikon kiépülését és a lexikális hozzáférési folyamatokat, és nem utolsósorban a beszédfeldolgozás elsajátításának mérföldköveit.

Hámori Ágnes és Bóna Judit 5 és 9 éves korú magyar gyermekek beszélőváltás-időzítési jellemzőit vizsgálja felnőttel való társalgásuk során. A szerzők társalgáselemzési és fonetikai módszerek segítségével a gyermekek beszélőváltási időtartamait (FTO) és a fordulók hosszát, valamint a gyermekek átlagos beszédtempóját és ezek összefüggéseit elemezték.

Hattyár Helga a jelnyelv-elsajátítás folyamatát ismerteti. Különös figyelmet szentel a magyar jelnyelv elsajátításának. Bemutatja a siketség eltérő elméleti és gyakorlati megközelítéseit, a jelnyelvek elsajátításának általános jellegzetességeit, módozatait és folyamatait, a bimodális kétnyelvűség jellegzetességeit és hatásait a nyelvfejlődésre, valamint a halló családba születő siket gyermekek eltérő igényeit és nehézségeit.

Kas Bence és Lőrík József a KOFA (Kommunikatív Fejlődési Adattár) segítségével végzett vizsgálat eredményei alapján a korai magyar nyelvfejlődés jellegzetességeit mutatják be. A nagy mintán alkalmazott eszköz szülői beszámolón alapuló módszertan alapján kidolgozott magyar nyelvre adaptált mérőeszköz. A 15-32 hónapos magyar gyermekek nyelvfejlődésére vonatkozó vizsgálat eredményei az expresszív szókincs, egyes morfológiai elemek, valamint a mondatalkotás elsajátításáról és fejlődéséről nyújtanak tájékoztatást.

Ladányi Enikő a ritmikai és nyelvi képességek kapcsolatát tipikus és atipikus nyelvi fejlődésű gyerekeknél vizsgáló tanulmányokat mutat be azzal a céllal, hogy felhívja a logopédus szakemberek figyelmét a gyenge ritmikai képesség jelentős kockázati tényező voltára, és a zenei ritmikai képességek fejlesztésének szükségességére a logopédiai gyakorlatban. Ugyanakkor javaslatot tesz az óvodás korban történő ritmikai képességek szűrésére és mérésére a későbbi nyelvi elmaradás és nehézségek kiküszöbölése érdekében.

Murányi Sarolta a narratív készség elsajátítását és fejlődési folyamatát mutatja be kétéves kortól kisiskolás korig. Ismerteti a téma elméleti hátterét és a hozzá kapcsolódó fogalomkört, az elsajátítást befolyásoló tényezőket, az otthoni nyelvhasználat hatását és az iskolai sikerességre gyakorolt befolyását. Külön fejezetet szentel a narratív készség elsajátítására vonatkozó magyar nyelvű kutatási eredmények bemutatására. Ebben a készségfejlődés folyamatában körvonalazódó mérföldkövekről kapunk tájékoztatást.

Neuberger Tilda a fonológiai fejlődéssel foglalkozik. Mássalhangzók hosszúsági kontrasztjait vizsgálja 7-8 éves gyermekek spontán beszédében. A szingleton és gemináta mássalhangzók időtartamának elemzése mellett a szomszédos magánhangzók időviszonyait is elemzi zöngétlen explozívák és nazálisok esetében. A vizsgálat eredményei új adatokkal szolgálnak a magyar anyanyelvű gyermekek fonológiai fejlődésének megismeréséhez.

Pléh Csaba az atipikus nyelvi fejlődés folyamatának egyik sajátos megnyilvánulási formájára irányítja a figyelmet. A specifikus nyelvi zavar/nyelvfejlődési zavar elméleti hátterét diakronikus nézőpontból mutatja be, a jelenséget kiváltó tényezőkről való tudományos gondolkodás folyamatába enged betekintést. Felvázolja a folyamat főbb állomásait, és feltárja az elért kutatási eredmények hatására bekövetkezett terminológiai változ(tat)ás okait.

Svindt Veronika a rendelkezésre álló fMRI-vel készült kutatások eredményei alapján a pragmatikai képesség működéséhez szükséges neurális hálózatok fejlődésének és érésének folyamatát tekinti át. Képet kapunk a pragmatikai fejlődést meghatározó hálózatok (a nyelvi hálózat, a szalienciahálózat, a tudatelméleti hálózat és a központi végrehajtó hálózat) strukturális és funkcionális fejlődéséről csecsemőkortól kamaszkorig, a folyamat mérföldköveiről, és arról, hogy ezeknek a hálózatoknak a fejlődése hogyan teszi lehetővé a gyerekek pragmatikai megértését.

Szeberényi Noémi és Gervain Judit a születés előtti nyelvi tapasztalatok hatását vizsgálják a nyelvelsajátítás folyamatára. A szerzők áttekintik a nyelvelsajátítás legkorábbi szakaszának empirikus vizsgálataiból származó eredményeket, és arra a kérdésre keresik a választ, hogy a korai nyelvi tapasztalatnak milyen hatása van a későbbi nyelvelsajátításra. Bár a nyelvelsajátítás folyamatának kezdeti szakasza az anyanyelvfejlődés kutatásában korábban viszonylag kevesebb figyelmet kapott, a vizsgálati eredmények áttekintése során az derül ki, hogy a magzatkori nyelvi tapasztalatnak fontosabb szerepe van, mint azt korábban gondolták. A korai nyelvi tapasztalat legfőbb eleme ugyanis a prozódia, amely a nyelvi fejlődés egyik fontos motorja. Ez alapján a nyelvi fejlődés új modelljét dolgozzák ki, amely a korai nyelvi tapasztalatok hatásait is figyelembe veszi.

A kötetbe foglalt szakirodalom hasznos olvasmány nemcsak a kutatók és egyetemi hallgatók, hanem a gyakorlati szakemberek, az anyanyelvi fejlődés iránt érdeklődők, illetve szülők számára is, mivel a téma iránt érdeklődő szakértő- és laikus olvasóközönség számára egyaránt naprakész információval szolgál. A tanulmánykötet a Kiadó okoskönyvtárában, a mersz.hu oldalon online is elérhető.

BALÁZS BORÓKA

Eötvös Loránd Tudományegyetem

borokabalazs@gmail.com

<https://orcid.org/0009-0000-2820-576X>

Domonkosi Ágnes:
**A megszólítás gyakorlatai. Tegezés, magázás, szólítás a társas
viszonyok alakításában**

(Budapest: Szépirodalmi Figyelő Alapítvány. Gondolat-Jel 9. 2024. 422 o.)

A könyv szerzője, Domonkosi Ágnes közel 30 éve foglalkozik a társas kapcsolatok különböző nyelvi megnyilatkozásival, kiemelten a megszólítás különböző szinterein megjelenő formáival. Doktori értekezését is a témából írta, amely *Megszólítások és beszédpartnerre utaló elemek nyelvhasználatunkban* címmel 2002-ben jelent meg Debreceni Egyetemi Kiadóban. A témában a szerzőnek számos tanulmánya látott napvilágot különböző szakfolyóiratokban, konferenciakötetekben is. Most bemutatandó könyve a legújabb kutatásainak, a 2017–2020 közötti időszak eredményeinek összegzése. Munkáját a Bolyai János kutatási ösztöndíj támogatásával végezte. A könyv anyaga több szempontból is újat hoz a korábbi kutatásaihoz viszonyítva: „...elméleti szinten a variabilitás dinamizmusként való értelmezése révén, módszertani szempontból a valós diskurzusok megszólításainak működését leíró elemzések, illetve a metadiskurzusok kvalitatív tartalomelemzésének köszönhetően; emellett pedig az eltelt idő miatt új, változásokat mutató tendenciák felvázolásának lehetőségéből adódóan is [...] A kötet funkcionális kognitív nyelvészeti keretben tárgyalja a megszólítás elméleti és empirikus aspektusait” (Domonkosi, 2024: 8).

A kötet anyagát, a *Bevezetést* követően, amely a megszólításkutatásban alkalmazott alapfogalmakat és a megszólítások kutatásában alkalmazott módszerek ismertetését tartalmazza, két nagy egységre tagolta a szerző. Az első, a *Szempontok a megszólítási gyakorlatok értelmezéséhez* 12 alfejezetből áll, amelyekben hangsúlyozottan az elmélet áll a középpontban: „... egyfelől a megszólítások témakörének központi kérdéseit helyezi előtérbe, tárgyalva többek között a megszólítás szerepét a személyközi viszonyok alakításában, a nézőpont érvényesülését a megszólítások műveletében, a személynevek és névtípusok szerepét a megszólítások műveletében, miközben összegző szándékkal a korábbi empirikus adatok értelmezését is tartalmazza” (Domonkosi, 2024: 12–13). A fejezet főbb témái: a megszólítások és köszönések; a nemtegezés változatainak szociokulturális értéke; szimmetria és aszimmetria a megszólítási gyakorlatokban; megszólítások és nemi szerepek; metaforikus és metonimikus motivációk a megszólítási gyakorlatokban; személynév és megszólítás.

A második fejezet, *A megszólítás gyakorlatai: helyzetek, szokások, értelmezések* című fejezet 13 alfejezetből áll. A teljesség igénye nélkül a következő témaköröket emelem ki közülük: a gyógyító-beteg viszony megszólítási változatai; megszólítási gyakorlatok a rádiózás nyelvhasználatában; megszólítási gyakorlatok az oktató-hallgató viszonyban; a megszólítás kérdései a nyelvi tanácsadói gyakorlatban.

A szerző által elemzett minden témakör rendkívül izgalmas, a gyakorlati élet különféle területein, a szóbeli és írásbeli megnyilvánulásokban előforduló kapcsolatteremtési gyakorlatot mutatja be és elemzi. A vizsgált korpuszok számos valós diskurzust, valamint metadiskurzusok elemzését tartalmazzák. „Mivel a nyelvi kapcsolattartás, a megszólítások kérdése a társadalom, a közösségek viselkedési normáinak fenntartásában is lényegi szerepet játszik, az elemző részek sorát a nyelvi tanácsadásról szóló, a normaközvetítés lehetséges módjait is problematizáló fejezet zárja” (Domonkosi, 2024: 7).

Ismertetésemben a megszólítási gyakorlatok két használati területét emelem ki, amelyek egyrészt mindannyiunk hétköznapi kommunikációjában is szerepet játszhatnak, másrészt az oktatási gyakorlatban is megfigyelhetők.

Az egészségügyben, különösen a kórházi osztályokon folyó megszólítási formákat sokan érzékelték már közülünk, sőt esetleg kifogásolhatták is. A szerző ebben a fejezetben a tegezés-nemtegezés, az esetleges nominális megszólítási formák általánosan is jellemző használatát mutatja be. Az orvosok és az ápoló személyzet, valamint a betegek különböző módon alkalmazzák a megszólítási formákat, amelyek a kommunikációban részt vevők társadalmi helyzetétől, műveltségi szintjétől és korától is függenek. A vizsgált közegben az orvos a magázó formát természetesnek tartja. A *maga* és az *ön* névmások használata azonban már más megítélés alá esik az orvos és a betegek részéről. Az orvos megítélésében a *maga* megszólítás közvetlenebbnek, szolidárisabbnak hat, mint az *ön*, „... illetve a területi kötöttségű szokásrend egyaránt szerepet játszhat a nemtegező névmásokhoz való viszonyulásban. A tetszikelő megoldásokat (a főorvos asszony) főként a nála idősebb, vagy valamilyen okból fokozottan tiszteletreméltónak tartott (például régi tanárok), illetve rangosabb betegekhez fordulva használja, nemcsak a nőkhöz, hanem a férfiakhoz is, de önreflexiója szerint „esete válogatja”, és a forma választása nem is mindig tudatos” (Domonkosi, 2024: 182). Nézzük meg a nem orvos egészségügyi szakszemélyzet nominális megszólítási változatainak elemzését! Ismeretlen 65 éves férfi beteg megszólítása (a fiatalokat általában tegezik) a kérdőívek tanúsága szerint: Szabó úr (semleges), Szabó bácsi (42%), János bácsi (16%), János (13%), Kedves János (2%), Szabó János bácsi (5%). A fiatalabb nőbeteget a keresztnévükön vagy annak becézett formájával, az idősebb betegeket a *néni* szóval kiegészítve szólítják meg: Borbála (60%), Borika (20%), Kovács néni (10%), Borika néni (5%), Kovácsné drága (5%). A férfiaknál megjelenő *úr*-nak megfelelő elem nem szerepel. Ez azt mutatja, hogy a magyar nyelvhasználati gyakorlatban nem alakult ki normaként elfogadott változata, mint pl. az angolszászoknál. A betegek az orvost *főorvosnő*-nek vagy *főorvos asszony*-nak szólítják a vizsgált osztályon. Az adatok alapján a nővérek, ápolók általában a nemtegező magázó vagy tetszikelő változatot használják, kivéve, ha jó ismerős vagy rokon a beteg.

A kutatásban szerepelt az egészségügy megszólítási formáinak és problémáinak vizsgálata az internetes médiadiskurzusban is. „Az összegyűjtött párbeszédekben

szereplő reflexiók alapján a betegek megszólításában alapvetően a bizalmasnak tartott, így akár az orvosi kommunikáció közelítő stratégiájának részeként is értelmezhető megszólítási módok, azaz a keresztnévi formák, *néni*, *bácsi* és a különböző családi megszólítások (pl. *mama*, *papa*, *anyuka*) kiterjesztett használata tűnik igen ellentmondásosnak, akár erőteljes érzelmi reakciókat is kiváltva” (Domonkosi, 2024: 198).

A megszólításformák oktatásban való használata különböző témaköreiből a tegezés-nemtegezés formáit mutatom be röviden. „A hallgatók sajátos társas helyzetben vannak a felsőoktatási intézményeken belüli hierarchiát, az oktatókhoz való viszonyukat tekintve. Az oktató-hallgató kapcsolat ugyanis részben szimmetrikus viszonyt jelent, hiszen leendő szakemberek, kollégák képzését végzik az intézmények, részben azonban aszimmetrikus is a hierarchikus szerveződés miatt” (Domonkosi, 2024: 313). A kutatásban többféle módszert alkalmazott a szerző: kérdőívek, oktató-hallgató levelezés, az oktatókkal készített interjúk, tanulói csoportokkal folytatott beszélgetések elemzése.

A témához tartozó vizsgálat két egyetemi közösség, az egri EKE és a (szlovák) komáromi SJE megszólítási szokásrendjét vizsgálta a magyar szakos hallgatói közösségekben. Az adatok elemzésével kapott eredmények azt mutatják, hogy az egri közösségben a tanár-diák kommunikációban a kölcsönös nemtegező formák vannak többségben. A nemtegezés változatai azonban igen összetett képet mutatnak. „A hallgatók tipikusan önözve fordulnak az oktatókhoz, de ahogy az interjúk is megerősítik, a legmeghatározóbb gyakorlat a névmáskerülés, azaz a 3. személyű formák nominális megszólítással (*tanár úr*, *tanárnő*) kiegészülő változatai. Az oktatók életkora is meghatározza a választott nemtegező formát. A vizsgálat eredményei azt mutatják, hogy az oktatók több mint a fele 40 év alatti, az ő névmáshasználatukban gyakrabban fordul elő az *ön*, mint az idősebbek névmáshasználatában. Álljon itt példaként egy középkorú női oktatóval készült interjúrészlet: „Igyekszem elkerülni a magát meg az önt, mindig megkérdezem a hallgató nevét, sőt még azt is az elején, hogy milyen megszólításnak örülne, de szerintem egyébként keverem, én jobban szeretem az önt, de amúgy meg azt mondják, hogy a maga kedvesebb” (Idézi Domonkosi, 2024: 317).

A komáromi közösségben is többségben a nemtegező megszólítási formák a megszokottak. „A nemtegezés változatai, azaz az önözés, magázás, a tetszikelés és a névmáskerülés ... oktatók és hallgatók általi használatának eltérő aránya a szerepviszonyokban érvényesülő hierarchiát jelzik” (Domonkosi, 2024: 320). A szlovákiai magyar közösségben a magázás az elfogadottabb. Amíg azonban Egerben az idősebb oktatók részéről elfogadott, Komáromban a hallgatók is nagy arányban használják a *maga* névmást, mivel ezt közvetlenebbnek tartják, és nem sértőnek a tanárral szemben. „Az eltérés okai között meghatározónak tűnik, hogy ebben a régióban sokkal kevésbé társul a *maga* névmáshoz az alá-fölé rendelt viszonyra utaló társas jelentés, kölcsönösen nemtegező viszonyban jóval

semlegesebb értékű formának számít, mint a magyarországi használatban” (Domonkosi, 2024: 322).

A könyv minden fejezete jól alkalmazható a nyelvi kultúra fejlesztésében az oktatás minden szintjén. Sajnos nagyon időszerű is a kérdés, mivel nyelvhasználatunkban sok a bizonytalanság, és ez nemcsak a megszólítási gyakorlatra igaz, hanem a társadalom különböző csoportjaiban megnyilvánuló nyelvi udvariasság szintjén is megnyilvánul: a közszereplők beszédében, a médianyelvben stb.

Az igényes munka rendkívül alapos elméleti anyagát és gyakorlati elemzéseit jól kiegészíti a könyv végén található gazdag felhasznált irodalom, valamint a kutatásban alkalmazott kérdőívek. Jó szívvel ajánlom a könyvet minden szakembernek és nem szakembernek egyaránt, aki szeretné igényesen alakítani nyelvi érintkezéseit a különböző társas színtereken.

A könyv jól használható segédkönyvként az oktatás minden szintjén, kiemelten az egyetemi magyar szakos tanárképzésben és a doktori képzésben is, például a szociolingvisztika, a nyelvi kultúra oktatásában.

LŐRINCZ JULIANNA

Eszterházy Károly Katolikus Egyetem Magyar Nyelvészeti Tanszék

Selye János Egyetem Magyar Nyelv és Irodalom Tanszék

jel2ster@gmail.com

<https://orcid.org/0000-002-1748-703>

**Terminológiai csütörtök. XI. Szakmai nap:
Terminológia, fordítás, szakszövegírás
Lehetőségek és elvárások a munkaerőpiacon**

(Károli Gáspár Református Egyetem, online, 2024. február 8.)

A címben is szereplő terminológiai csütörtök immár a tizenegyedik a szakmai napok sorában. A terminológiai csütörtököt már 2014 óta minden évben február elején szervezi a Károli Gáspár Református Egyetem (KRE) Magyar Nyelvtudományi Tanszéke, a Terminológiai és Kommunikációs Kutatócsoportja (TERMIK), és a Magyar Nyelv Terminológiai Tanácsa (MaTT). A korábbi években a terminológiai mesterszakhoz kapcsolódva szervezték, az utóbbi két évben pedig a KRE által meghirdetett *Magyar irodalom és nyelvészet mesterszak* 50 kredites *alkalmazott nyelvészet* nevű specializációját mutatták be és népszerűsítették az előadók. A workshopot Fóris Ágota egyetemi tanár, a TERMIK vezetője, a MaTT elnöke nyitotta meg. A programban összesen öt előadást hallhatunk, amelyek témája a munkaerőpiaci igények köré szerveződött. A levezető elnök Sólyom Réka egyetemi docens volt.

Az első előadást Bölcskei Andrea és Fóris Ágota tartották *Szakszövegírás, terminológia, tartalommenedzsment, alkalmazott nyelvészeti specializáció a magyar mesterszakon* címmel. Az előadást Fóris Ágota kezdte a terminológiai mesterképzési szakról, amely idővel sajnos megszűnt, ugyanis az EMMI nem vette fel a felsőoktatási szakok listájába, emiatt is hozták létre az új specializációt. A fókusz a nyelvészen és alkalmazott nyelvészen van, melyben a terminológia, a szakszövegírás és a tartalommenedzsment a fő téma. Megtudhattuk azt is, hogy a végzett hallgatók képesek voltak elhelyezkedni ebben a szakmában, itthon és külföldön is, például az Akadémiai Kiadónál, de vannak, akik szabadúszó fordítók lettek vagy multinacionális cégeknél dolgoznak, esetleg PhD fokozatot is szereztek. Az előadó azt is elmondta, hogy gyakran szerveznek kerekasztal-beszélgetéseket arról, hogy a munkaerőpiacon hol és hogyan lehet elhelyezkedni, továbbá rendszeresen vizsgálják, hogy milyen tudásuk, képzettségük legyen a továbbiakban a hallgatóknak. Az előadás további részét Bölcskei Andrea tanszékvezető, intézetvezető tartotta magáról a szakról. Elsősorban a szakszövegírásról és annak feladatköréről, hasznosságáról tudhattunk meg érdekességeket. A feladatkörnek megfelelően azokat a készségeket és képességeket is felsorakoztatta, amelyekre egy szakszövegírónak szüksége van: ilyenek például a jó nyelvérzék, az elemző gondolkodás, a célközönség iránti empátia stb. A következőkben azokról a felsőoktatási képzések között jelen levő szakokról tudhatunk meg többet, amelyek keretében szakszövegírókat képeznek: itt a Limerick Egyetem mellett a Berkeley Egyetemet emelte ki és mutatta be részletesebben. Ezt követően megismerhettük a szakszövegírók szervezetének ajánlását a felsőoktatási képzésekről az intézmények számára. Hatféle irányt neveztek meg a megfelelő szakismeretek

megszerzésére, így ebbe tud bekapcsolódni a magyar nyelv és irodalom mesterképzési szak, ahol szakszövegírás, terminológia és tartalommenedzsment specializációt terveztek. Bővebben bemutatásra kerültek az alkalmazott nyelvészet specializáció témakörei (többek között pragmatika, szókészlet, névtan, szakszövegírás és dokumentáció stb.). Az utolsó részben az előadó a mesterséges intelligencia és a nagy nyelvi modellek lehetséges szerepét tekinthette át, majd a mintatantervet ismertette.

A következő előadást Sermann Eszter, az SZTE egyetemi adjunktusa tartotta *Terminológiai kompetencia elméleti és gyakorlati megközelítésben* címmel. Az előadás első felében a terminológus feladatköreit mutatta be, majd ezt folytatva Sermann egy táblázat bemutatásával, a TermCoord előírásait is figyelembe véve szemléltette, hogy egy terminológus mihez ért, illetve mihez is kell értenie. Sermann az előadásban fontosnak tartotta megemlíteni a fordítók számára szükséges terminológiai készségeket is. Ezután rátért a vizsgálat leírására, és elmondta, hogy kutatását a közösségi hálózat internetes felületén végezte el a Magyarországra vonatkozó álláshirdetések között. A továbbiakban a terminológiakezelést igénylő munkaköröket ismertette, majd betekintést nyerhettünk a keresési eredményekbe is, amelyeket 2024 januárjában végzett el.

Faludi Andrea (LEG Zrt.) *Terminológiai adatbázisok alkalmazási lehetőségei a szakszövegírásban és a szakfordításban* címmel tartotta előadását. A célja a különböző terminológiai adatbázisok felhasználásának bemutatása volt. Elsősorban a terminológiai adatbázisokról tudtunk meg olyan információkat, mint a funkciójuk, céljuk és az alapvető adatkategóriáik, majd azt is, hogy milyen hibákat tudunk elkerülni a fordítás során, amikor ezeket a terminológiai adatbázisokat használjuk. Faludi Andrea emellett kitért a címben is megjelenő szakszövegírásra és a fordítás előkészítésére, és különböző szemléltető képekkel mutatta be, hogy hogyan is kapcsolódnak ezek a terminológiai adatbázisok használatához. A fordításhoz kapcsolódóan az előadó a gépi fordítás során használt adatbázisokat is ismertette, több különböző táblázattal illusztrálva a találatokat.

A következő előadó Papp Eszter volt, aki *Soknyelvű terminológiai munka humanitárius területen* címmel tartott előadást. Papp Eszter a CLEAR Global nevű nemzetközi cégnél dolgozik, előadásának első részében bemutatta magát a szervezetet, a náluk dolgozó önkénteseket, alkalmazottakat és feladataikat, majd áttért a humanitárius terminológia jellemzőire, mint amilyen a magas szintű absztrakció, az angol munkanyelv, érzékeny témák stb. Papp az általa felsorolt hat jellemző közül az *érzékeny témákat* külön kiemelte és részletesen bemutatta. Ezt követően arról a fordításorientált terminológiai munkáról beszélt, amelyet önkéntesek segítségével végeznek. Szintén a terminológiai munkához kapcsolódóan mutatta be azt a nagyobb lélegzetű doménorientált munkát, amely a szervezetnél folyik. A munkák során született eredmények az ún. Glossary app-be kerülnek, azon belül is a CLEAR glosszáriumokba, amelyből tizennégy darab

érhető el nyilvánosan. Ezt követően a PSEA glosszáriumot ismertette részletesebben, majd a mesterséges intelligenciának ezekben a munkákban lehetséges használatáról beszélt.

Az utolsó előadást Tamás Dóra Mária és Magas Ágnes terminológusok tartották meg *Betekintés az OFFI terminológiai műhelyébe* címmel. Az előadást Tamás Dóra Mária egy áttekintéssel indította, majd röviden bemutatta az OFFI Zrt. tevékenységeit. Ezek után az OFFI terminológiai projektjét ismertette (terminológiai adatbázis építése), és a terminológiai témájú kiadványokat is bemutatta. Magas Ágnes a terminológiai munka típusairól beszélt, amelyet néhány gyakorlati példán keresztül mutatott be.

A programot Sólyom Réka zárta, és invitálta a hallgatóságot kérdések feltevésére, majd bejelentette a tudományos programot követő könyvbemutatót. Ennek során Prószéky Gábor, a HUN-REN Nyelvtudományi Kutatóközpont főigazgatója, a kötet főszerkesztője mutatta be a 2023-ban megjelent *A magyar terminológiastratégia kialakítása Zöld könyv* című könyvet (doi: <https://doi.org/10.18135/term.2023>). Részletesen ismertette a kiadásának célját, bemutatta a kötet íróit, szerkesztőit, ezt követően a könyv szerkezetét, amelyet négy fejezetre osztottak: I. A terminológiastratégia általános kérdései, II. A terminológiastratégia speciális kérdései, III. Szabványok és helyesírás, IV. Bibliográfia. Prószéky az előadásában ezeket a fejezeteket részletesen is bemutatta. A rendezvény Fóris Ágota zárszavával ért véget.

LUTA ADRIENN
Károli Gáspár Református Egyetem
lutaadri888@gmail.com

**Tamás Dóra Mária – Szoták Szilvia (szerk.):
Terminológiastratégia és jogi alapterminusok a szomszédos
országok nyelvén**

(Budapest: OFFI Zrt. 2021. 384. o.)

<https://www.offi.hu/offi-akademia/kiadvanyok/terminologiastrategia-es-jogi-alapterminusok-a-szomszedos-oroszagok-nyelven>

A *Terminológiastratégia és jogi alapterminusok a szomszédos országok nyelvén* című könyv célja, hogy 193 jogi kulcsterminus magyar és hat szomszédos országbeli megfelelőjét bemutatva segítse a jogi szakfordítást, a fordítóképzést, és hozzájáruljon a magyar jogi nyelv megőrzéséhez a digitális térben. A kiadvány Németh Gabriella által írt sorozatszerkesztői előszava a jogi nyelv és terminológia folyamatos változására fókuszál. A szöveg rámutat, hogy a jogi nyelv, bár alapvetően hagyománytisztelő, alkalmazkodik a modern nyelvi és technológiai változásokhoz. A jogi terminusok változása a mögöttük álló társadalmi, gazdasági és politikai változásokat is tükrözi, ami kihívást jelent a fordítók és lektorok számára.

A szerkesztői előszóból, melyet Szoták Szilvia és Tamás Dóra Mária írt, megtudhatjuk, hogy a kötet kiemeli a jogi terminusok fontosságát a magyar nyelv versenyképessége szempontjából. Az OFFI Zrt. már régóta végez terminológiai munkát, 2016 óta tudatosan építi a közigazgatási és jogi terminusok adatbázisát. Az előszó történeti kontextusba helyezi az OFFI feladatait, bemutatva a Központi Fordító Osztály 1869-es megalakulását és a kisebbségi nyelvek szerepét. A kötet célja a szomszédos országok terminológiapolitikájának és a magyar kisebbségek helyzetének bemutatása, a fordítói munka terminológiai támogatása és az egységes minőség megteremtése. A könyv két részből áll: az elméleti rész tanulmányokat tartalmaz a terminológiastratégiáról és az igazságügyi nyelvstratégiáról, míg a második rész a szomszédos országok nyelvén a leggyakrabban előforduló jogi terminusokat írja le.

Az első tanulmány, *Az igazságügyi közfeladat – ellátás nyelvstratégiai kérdései a jog és nyelv vonatkozásában*, az igazságügyi nyelvhasználat nyelvstratégiai kérdéseit vizsgálja. Németh Gabriella és Szoták Szilvia rámutatnak, hogy az igazságügy széleskörű és összetett rendszer, amely áthatja a közigazgatást és a társadalmat, ami viszont megnehezíti az egységes nyelvi politika kialakítását. Javasolják igazságügyi nyelvstratégia kidolgozását, melynek célja a közérthető jogi nyelvhasználat elősegítése, a terminológia egységesítése és a nyelvi tudatosság növelése. Az OFFI, mint az igazságügy fontos szereplője, már lépéseket tett a terminológia egységesítése irányába egy belső elérhető terminológiai adatbázis, a IUSTerm létrehozásával. A stratégia fontos eleme lenne a digitális tér nyelvi hatásainak vizsgálata is, hiszen az online térben közzétett jogi tartalmak elérhetősége és érthetősége kulcsfontosságú a jogbiztonság és a demokratikus joggyakorlás szempontjából. A szerzők hangsúlyozzák az

interdiszciplináris együttműködés szükségességét a stratégia kialakításában, bevonva a jogászokat, nyelvészeket és az igazságügy egyéb szereplőit.

A második írás Fóris Ágota *Bevezető gondolatok a terminológiastratégiáról* című tanulmánya. A szerző a magyar terminológiastratégia fontosságát hangsúlyozza, különös tekintettel a szomszédos országokban élő magyar nyelvközösségekre. Rámutat, hogy a terminológia egységesítése és harmonizációja elengedhetetlen a magyar nyelv versenyképességének megőrzéséhez és a határon túli magyarok anyanyelvhasználatának támogatásához. A tanulmány bemutatja az OFFI szerepét a terminológiai munkában, és kiemeli a IUSTerm jogi és közigazgatási terminológiai adatbázis létrehozását. Hangsúlyozza egy nemzeti terminológiai adatbázis létrehozásának fontosságát, amely a magyar nyelvű terminológia mellett angolul és a szomszédos országok hivatalos nyelvein is tartalmazná a terminusokat.

A következő írás Tamás Dóra Mária *Jogi terminológia a terminológus szemével* című tanulmányából származik, és az OFFI terminológiai tevékenységét, valamint a *Terminológiastratégia és jogi alapterminusok a szomszédos országok nyelvén* című kötet célkitűzéseit mutatja be. A tanulmány rávilágít a jogi terminológia sajátos jellemzőire, mint például a naprakészség, az erős verbalitás, a rendszerbe ágyazottság és az országspecifikusság. A szerző bemutatja a teljes ekvivalencia, a részleges ekvivalencia és az ekvivalencia hiányának eseteit, valamint a funkcionális ekvivalens és a terminusjelölt fogalmát. A tanulmány bemutatja a közérthetőség fontosságát a jogi szaknyelvben, de kiemeli, hogy az egyszerűsítés nem mehet a pontosság rovására. A szerző példákon keresztül szemlélteti a különböző fogalmak pontos meghatározásának és megfelelő fordításának fontosságát. Rámutat, hogy a terminológiapolitika a nyelvpolitika része, és célja a magyar nyelv versenyképességének megőrzése és a nyelvi jogok érvényesítése.

A negyedik tanulmány Benő Attila, Kokoly Zsolt és Benő Lóránd *A román és a magyar jogi terminológia helyzete Romániában – szaknyelvi problémák és terminológiatervezési feladatok* című írása, amely bemutatja a romániai magyar jogi szaknyelv történeti fejlődését, kiemelve az 1989 utáni revitalizációs törekvéseket. Részletesen ismertetik a Szabó T. Attila Nyelvi Intézet szerepét a kétnyelvű szakszótárak készítésében és a terminológiai adatbázisok fejlesztésében. Rámutatnak a magyar és román jogrendszerek eltéréseiből adódó terminológiai nehézségekre, különösen az elvontabb fogalmak, intézménynevek és tisztségnevek magyarításának problematikájára. A szerzők kiemelik a Sapientia Erdélyi Magyar Tudományegyetem szerepét a magyar nyelvű jogászképzés megteremtésében és a jogszabályok magyar nyelvű fordításában. A szerzők következtetése az, hogy a romániai magyar jogi szaknyelv a magyar jogi szaknyelv egy sajátos, a román jogrendszerhez igazodó változata, amelynek fejlesztése elengedhetetlen a romániai magyarok jogbiztonsága és anyanyelvhasználatának megerősítése szempontjából.

A következő mű Szabó Mihály Gizella *A szlovák nyelvű és a magyar jogi terminológia helyzete Szlovákiában* című tanulmánya, mely elemzi a szlovák és a szlovákiai magyar jogi terminológia történeti fejlődését, jelenlegi helyzetét és a terminológiastratégia fontosságát. Részletesen bemutatja a szlovák terminológia történetét, kiemelve a cseh, a magyar és az angol nyelv hatását, valamint a különböző történelmi korszakok (Csehszlovákia, a II. világháború utáni időszak, rendszerváltás) terminológiára gyakorolt hatását. Kitér a szlovák terminológiapolitikára, elemezve a nyelvi törvényeket, kormányzati dokumentumokat és terminológiai adatbázisokat (*STD, Tezaurus slovenského práva*). A szerző rámutat a terminológiai képzés hiányosságaira Szlovákiában és elemzi a szlovák terminológiai hálózat szerepét. A tanulmány a szlovák jogi terminológia sajátosságait vizsgálva kiemeli a motiváltság, a rendszerszerűség, az egyértelműség és a pontosság fontosságát. Megállapítja, hogy a szlovákiai magyar jogi szaknyelv a magyar jogi szaknyelv sajátos változata, melynek fejlesztése elengedhetetlen a szlovákiai magyarok jogbiztonsága és anyanyelvhasználatának megerősítése szempontjából.

Beretka Katinka és Lucil Emil *A szerb nyelvű és a magyar jogi terminológia helyzete, avagy az ország, ahol előlátják a büdzsé-rebalanszt* című tanulmánya a szerbiai, különösen a Vajdaságban élő magyarok jogi terminológiájának helyzetét és a terminológiastratégia jövőbeli lehetőségeit tárgyalja. Kiemelik, hogy bár a hivatalos nyelvhasználatról szóló törvény elméletileg biztosítja a magyar nyelv használatát a bírósági és közigazgatási eljárásokban, a gyakorlatban jelentős hiányosságok mutatkoznak. A tanulmány bemutatja a szerb és a magyar jogi terminológia közötti különbségeket, a fordítás során felmerülő nehézségeket, és rámutat a magyar nyelvű jogászképzés, a fordítóképzés, a szaknyelvi adatbázisok és a terminológia-egységesítő tevékenység fontosságára. A tanulmány párhuzamot von a szerbiai helyzet és a korábban tárgyalt szlovákiai és romániai magyar jogi szaknyelv helyzete között, kiemelve, hogy mindhárom országban a magyar jogi terminológia a magyar jogi szaknyelv egy sajátos, a helyi jogrendszerhez és a nyelvi kontaktusokhoz alkalmazkodó változata.

A következő írás Lehocki-Samardžić Anna és Palizs-Tóth Hajnalka *A magyar nyelvű jogi terminológia helyzete Horvátországban* című tanulmánya. A szerzők bemutatják a horvátországi magyar jogi szaknyelv történeti hátterét, kiemelve a jugoszláv időszak hatását és a rendszerváltás utáni revitalizációs törekvéseket. Ismertetik a horvát terminológia standardizációjának folyamatát a *STRUNA* projekt és a Horvát Terminológiai Portál létrehozásával, összehasonlítva a magyar terminológia kialakulatlanságával. A tanulmány elemzi a magyar jogi terminológia jellegzetességeit Horvátországban. Bemutatja a terminológiai nehézségek típusait (például a horvát jogi rendszerre jellemző terminológiát), és hangsúlyozza a funkcionális ekvivalencia, a tükörfordítás és a parafrázis szerepét a terminológiai megfeleltetésben. Bemutatja a Glotta Nyelvi Intézet és az eszéki Magyar Nyelv és Irodalom Tanszék szerepét a terminológia fejlesztésében,

kiemelve az *Omniglosar* és a *Termini magyar–magyar szótár* jelentőségét. Párhuzamot von a horvátországi helyzet és a korábban tárgyalt szerbiai, szlovákiai és romániai magyar jogi szaknyelv helyzete között.

A nyolcadik rész Farkas-Kordonec Gabriella és Rácz Beáta tanulmánya, mely az ukrainai, különösen a kárpátaljai magyarok jogi terminológiájának helyzetét elemzi. Rámutatnak, hogy bár Ukrajna ratifikálta a regionális vagy kisebbségi nyelvek európai kartáját, a kisebbségi nyelvek, köztük a magyar nyelv használata a jogi területen jelentősen korlátozott, és hiányzik az egységes terminológiai adatbázis. Kiemelik, hogy nincsenek elektronikus terminológiai adatbázisok a kisebbségek nyelvén, a jogszabályok hivatalos fordítása pedig csak az orosz és az angol nyelvre korlátozódik. A meglévő szótárak hiányosak, és nem tartják a lépést a jogi nyelv változásaival. Bemutatják a fordítás során felmerülő nehézségeket, például a szabadságvesztés végrehajtási fokozatai terén. Párhuzamot vonnak az ukrainai helyzet és a korábban tárgyalt szerbiai, szlovákiai és horvátországi magyar jogi szaknyelv helyzete között.

A kötet utolsó fejezete a magyar jogi terminológia többnyelvű ekvivalenseit gyűjti össze, bemutatva a magyar jogrendszer sajátosságait és a fordítás során felmerülő nehézségeket. A tanulmány hangsúlyozza a kontextus és a terminológiastratégia fontosságát a pontos fordítás érdekében, kiemelve a funkcionális ekvivalensek és a terminusjelöltek használatát. A forrás számos jogterületet érint, például a büntetőjogot, családjogot, és cégjogot, illusztratív példákkal.

A kötet végén egy magyar jogi terminusgyűjtemény található, amely a terminusok definíciói mellett fordítási útmutatást is nyújt. A gyűjtemény hét fő jogterületet ölel fel: általános jog, bírósági szervezet, alkotmányjog, közigazgatási jog, büntetőjog, polgári jog és cégjog. A terminusok definíciói jogszabályokon, hivatalos honlapokon, tankönyveken és jogi lexikonokon alapulnak, és jogászok ellenőrizték. Legvégül egy tárgymutató következik a legfontosabb kifejezésekkel és kulcsszavakkal, illetve egy rövid bemutató a kötetben szereplő tanulmányok szerzőiről.

A *Terminológiastratégia és jogi alapterminusok a szomszédos országok nyelvén* című kötet átfogó képet nyújt a jogi szaknyelv és terminológia aktuális kérdéseiről. A kötet tanulmányai rámutatnak a közérthető jogalkotás és jogi kommunikáció fontosságára, a terminológia egységesítésének és harmonizációjának szükségességére a magyar nyelv versenyképességének megőrzése és a határon túli magyarok anyanyelvhasználatának támogatása érdekében.

ULLMANN CINTIA
Károli Gáspár Református Egyetem
ullmanncintia@gmail.com

Karen Stollznow: *Bitch: The Journey of a Word*

(Cambridge: Cambridge University Press, 2024. 200 p.)

Karen Stollznow ausztrál-amerikai nyelvész alkalmazott szemantikát és a szociolingvisztika egyéb területeit kutatja, különös tekintettel az előítéletekre és a diszkriminációra a nyelvhasználatban. Jelenleg az amerikai Griffith Egyetem kutatója, és társműsorvezetője a *Monster Talk* című díjnyertes, tudományos alapú podcastnak.

Bitch: The Journey of a Word című könyve arra vállalkozik, hogy nyelvészeti áttekintést nyújtson az angol *bitch* szó történetéről, a 'nőstény kutya' jelentéstől a feminizmusra gyakorolt mai hatásáig; feltárja a szó összetett és folyamatosan változó mai jelentéstartományát, és megmutassa, hogyan befolyásolja azt a beszélő neme, etnikai háttere és szexualitása; mindezt pedig széleskörű példatárra építve teszi, a történelmi dokumentumoktól kezdve a tévésorozatokon és szótárakon át a közösségi médiáig.

Stollznow nem nevezi meg, hogy kiket tekint a kötet célközönségének, de gyorsan egyértelművé válik, hogy a könyv nem csupán nyelvészek számára készült. Mindazok számára írta a szerző, akiket mélyebben érdekel a nyelv társadalmi szerepe és hatása, különösen a nem és a szexualitás metszéspontjában. Egyaránt szól nyelvészeknek és a nyelv iránt érdeklődőknek, feminista gondolkodóknak és aktivistáknak, valamint mindenkinek, aki szeretné jobban megérteni, hogyan válhat egyetlen szó – mint a *bitch* – kulturális és társadalmi viták középpontjává.

Az, hogy a kötet egy szélesebb olvasóközösséget kíván megszólítani, kitűnik szemléletes és közvetlen hangvételéből, amely az ismeretterjesztő, tudománynépszerűsítő művekhez hasonlóan teszi érdekessé a tudományos tényeket. A mindennapi életből és populáris kultúrából – például a *Simpson családból* – hozott példák és a szöveg hangvétele frissé és befogadhatóvá teszik a tartalmat, anélkül, hogy ez a mű tudományos értékét csökkentené. Stollznow értelmezést kínál olyan alapvető nyelvészeti terminusokhoz, mint például az *etimológia*, ismertetést nyújt az angol nyelv eredetéről, bebizonyítja, hogy a szavak jelentése idővel megváltozhat, és megmagyarázza, hogy a hozott óangol példák miért tűnnek úgy, „mintha idegen nyelven lennének írva”. A szöveg megértéséhez tehát nem szükségesek sem angol nyelvészeti, sem nyelvtörténeti előismeretek.

A kötet különösen izgalmas lehet nem angol anyanyelvű olvasók számára, hiszen feltár olyan nyelvi és kulturális rétegeket, amelyek a hétköznapi nyelvhasználat mögött húzódnak meg, és amelyek gyakran rejtve maradnak a tantermi nyelvtanulás során. Rávilágít arra, hogyan formálódik egy-egy szó jelentése a társadalmi kontextusok hatására, nemcsak nyelvi, hanem kulturális szemszögből is betekintést kínál a jelentésváltozások mögött álló folyamatokba.

A magyar szótárakban a *bitch* fordítása jellemzően 1. 'szuka' és 2. 'szajha', így például a HUN-REN SZTAKI angol–magyar szótárában is. A magyar beszélők

számára tehát valószínűleg nem egyértelmű, hogy az angol *bitch* szó mai jelentése, – amennyiben emberekre, nem pedig kutyákra értjük –, nem kizárólagosan, sőt nem is elsődlegesen ’prostituált’, hanem ’kellemetlen nő’ (18). Stollznaw jelentések sokaságát sorakoztatja fel, emellett pedig a szónak rengeteg olyan előfordulását mutatja be, amelyek az angol anyanyelvűek számára sem feltétlenül ismertek, mert már kivesztek a használatból, vagy mert egy adott területhez, vagy adott társadalmi réteghez köthetők. Így például a *stand bitch* kifejezés a történetiségben ’teát készít’ jelentéssel bírt, a *kitchen bitch* a karib-térségi angolban egyfajta ’hordozható lámpa’-t jelent, a Prodigy banda Smack My Bitch Up című számában található *bitch* pedig a karnak annak a vénájára utal, amelybe a heroint injekciózzák.

Stollznaw feltárja, miként alakult ki és változott meg a *bitch* szó jelentése az évszázadok során, valamint megvilágítja, milyen társadalmi erők – például nemi szerepek és hatalmi viszonyok – formálták a szó használatát és megítélését. Így ad a szóról igen széleskörű áttekintést.

Az A Bitchin’ Word (’Egy király szó’¹) bevezető fejezetben Stollznaw azzal indítja könyvét, hogy megosztja egy személyes tapasztalatát arról, amikor egy munkatársa a *bitch* szóval bántalmazta őt. Röviden kitér arra is, hogy miért válhatott a *bitch* ennyire népszerű sértéssé. Ezt követően felvázolja a könyv fent bemutatott témáját, főbb kérdéseit.

Az A Female Dog (’Egy nőstény kutya’) fejezetben a cím ellenére Stollznaw nem csupán a szó eredeti, ’szuka’ jelentéséről értekezik, – amely első írásbeli előfordulása 1000 körülre tehető (8) –, hanem bemutatja a szó teljes történetét, tehát a fejezet nagyobbik részét teszi ki az emberre vonatkoztatott *bitch*. Áttekintést ad arról, hogyan válhatott a *bitch in heat*, vagyis ’tüzelő szuka’ jelentésen keresztül előbb ’könnyűvérű nő’, ’prostituált’ jelentésűvé a szó, majd pedig valami még ennél is sértőbbé: ’kellemetlen, utálatos nő’ jelentésűvé. Ez a *bitch* már nem a nő szexualitását, hanem a személyiségét becsmérli, és bántóbb, mint a ’prostituált’ jelentésű *whore*, mert neki legalább fizetnek a szolgálataiért (19).

A fejezet a *bitch* szó nagyon sok különféle előfordulását ismerteti, a fent említett *stand bitch* és *kitchen bitch* kifejezéseket számtalan többi közül emeltem ki. A példák nem strukturált rendszer szerint követik egymást, tehát akár kronológiát vagy térbeliséget követve, de ez a megértést nem akadályozza, a szöveg folyása zavartalan és figyelemfenntartó.

A *bitch* mellett Stollznaw számos egyéb szót is bemutat, amelyek pejoráción mentek keresztül és nőkre alkalmazott sértésekké váltak, így például a *slut*, *hussy* és *tart* (’prostituált’), *mistress* (’szerető, kitarzott nő’), *shrew* és *wench* (’néember’). A fejezetet annak a vizsgálatával zárja, hogy milyen erők idézhetik elő ezeket a változásokat.

¹ A magyar fordítások tőlem származnak.

A *Feminist Bitch* ('Feminista ribanc') fejezet feminizmustörténeti áttekintést ad, emellett pedig azt követi végig, hogyan vált a *bitch* a mindennapos szóhasználat részévé. Stollznaw megállapítja, hogy a szó a nők szociális helyzetével együtt változott a 19. és 20. században, és az 1900-as évek közepétől terjedt el robbanásszerűen, válaszként a feminizmusra. A feministák egy része visszafoglalta, magáévá tette a szót, így például 1968-ban *The BITCH Manifestó*ban Jo Freeman, aki a *bitch*-et 'erős nő'-ként határozta meg.

A fejezet kitér arra, hogyan bántalmazták a *bitch* szóval a mai napig a nőket a politikában, és a politikán kívül, a magánéletben. A *bitch* gyakran megelőzi vagy kíséri a nők elleni erőszakot, úgy mint a szexuális zaklatást, családon belüli erőszakot vagy nemi erőszakot – számos tanulmány kimutatta, hogy a férfiak gyakran ezzel a szóval tárgyiasítják a partnerüket, mielőtt megütnék őt vagy más erőszakos tettet követnének el (54). A *bitch* tehát veszélyes szó, és az is nyitott kérdés marad, hogy lehetséges-e teljes mértékben visszafoglalni.

A *Son of a Bitch* ('Gazember') fejezet a *bitch* szó férfiakra történő alkalmazásával foglalkozik. Stollznaw kijelenti, hogy férfiak ellen majdnem olyan rég óta használják a *bitch* szót, mint nők ellen, viszont másképp: szemben a nőkkel, akik azért kapják ezt a sértést, mert túl erősek, a férfiak esetében a *bitch* gyengeségre utal, ami pedig egyben a nőkre is negatívan hat vissza. Akár férfira vagy nőre használják, a *bitch* szó azt sugallja, hogy az illető nem felel meg a társadalmi nemi elvárásoknak, vagyis nem a „megfelelő” módon teljesíti a neméhez társított szerepet.

A fejezet bemutatja a *son of a bitch* kifejezés történetét is, a férfi börtönök *prison bitch* kifejezését, illetve a *bitch*-et az LMBTQ-közösségben mint sértést és mint visszafoglalt szót. Emellett Stollznaw a férfiakkal kötődő nemi alapon megfogalmazott sértésekkel is foglalkozik. Röviden összeveti a férfiakra és a nőkre irányuló sértéseket, amelyeknek az összehasonlítása felhívja a figyelmet átfogó nyelvészeti kutatások szükségességére a témában.

A *The B Word* ('A *b* betűs szó') fejezet a *bitch* szónak a médiában való megjelenésével foglalkozik. A szerző bemutatja, ez a sokáig cenzúrázott szó hogyan vált egyre és egyre hétköznapiabbá, és mindeközben hogyan próbálták meg – sikertelenül – több ízben is betiltani. Ismerteti, a *bitch* hogyan szerepel a szótárakban, és hogyan érte el egy petíció, hogy az Oxford University Press *woman* ('nő') szócikkében ne szerepeljen a *bitch* szinonimaként.

Stollznaw hosszasan szemlélteti, hogyan vált a szó használata egyre gyakoribbá az újságokban, zenében, és filmekben egyaránt. A fejezet jelentős részét adja a szónak a jazzben és dirty blues-ban való gyökereinek, illetve a rapben és hiphopban elfoglalt helyének bemutatása. A hatalmas anyag feldolgozását a szerző alaposan és lényeglátóan végzi el, a példáival szemléletes és átlátható összefoglalását adja a *bitch* médiában való elterjedésének.

A *Basic Bitch* ('Klisé ribanc') fejezet azt a kérdést veti fel, hogyan is használják ma a *bitch* szót és valószínűleg hogyan fogják használni a jövőben. Ismerteti a szó

különböző variációit, – például *biznatch* és *biatch* –, illetve sok színes példát hoz különféle előfordulásaira. Ilyen többek között a fejezetcímekben is szereplő *basic bitch* ('klisé ribanc'), a *ride/sit bitch* ('motor hátsó ülésén utazik'), vagy a *go bitchcakes on someone* ('irracionális mértékben megharagszik valakire'). Stollznow felsorol továbbá dolgokat, amelyet *bitch*-ként lehet meghatározni – ide tartoznak a természeti katasztrófák, de az élet és a szerelem is.

Kitér a *bitch* szó jelnyelvi előfordulására, illetve a összegyűjti a megfelelőit különböző idegen nyelvekben, például kínaiban, lengyelben, héberben, arabban, zuluban, németben, újlatin nyelvekben és mesterséges nyelvekben: eszperantóban, a Trónok harca valériai nyelvében, az Avatar Na'vi nyelvében, klingonban.

A szó jövője nyitott kérdés marad, de a szerző a múltbeli viselkedés alapján kísérletet tesz a jövőbeli viselkedés előrejelzésére.

A Conclusion: A Final Bitch ('Konklúzió: Egy végső ribanc') fejezetben Stollznow a könyv összefoglalását adja, és kiemeli, hogy a legfontosabb a nők hátrányos helyzetének megváltoztatása – talán a *bitch* szó is visszafoglalhatóvá válik, ha ez sikerül.

A *Bitch: The Journey of a Word* nem csupán egy szó jelentéstörténeti vizsgálatát kínálja, hanem azt is megmutatja, hogyan formálja a nyelv a gondolkodásunkat, és hogyan tükrözi vissza a társadalomban mélyen rögzült előítéleteket, hatalmi viszonyokat és normákat. Átfogó képet ad arról, hogyan alakulnak a nemi szerepekhez és a nőkhöz kapcsolódó nyelvi attitűdök az angol nyelvben. A munka túlmutat a nyelvészet szűken vett keretein: társadalom- és kultúrtörténeti kontextusba helyezi a vizsgált kifejezést, rávilágítva arra, hogyan fonódik össze a nyelv a hatalom, az identitás, a szexualitás és a nemek közötti viszonyok alakulásával.

A kötet egyúttal érzékenyen reflektál a nyelvben jelenlévő nemi egyenlőtlenségekre, és bemutatja, hogy a nyelvi sértések és címkék miként járulnak hozzá ezeknek a struktúráknak a fenntartásához vagy épp megkérdőjelezéséhez. Tudományos alapossággal, de közérthető módon tárja fel a nyelv és a társadalom kölcsönhatásait, és tanulságos olvasmány mindazok számára, akiket érdekel a nyelv, a társadalmi nem, valamint ezek metszéspontjai.

Minden bizonnyal hasznos lenne a könyvben tárgyalt jelenségeket a magyar gyakorlattal összevetni. Egy ilyen összehasonlítás rávilágíthatna arra, milyen módon működik a nyelvi szexizmus különböző nyelvi és kulturális kontextusokban, és hogy milyen eszközökkel próbálják a beszélők ezeket újraértelmezni vagy épp megbélyegző erejüktől megfosztani.

SURÁNYI ANNA
Eötvös Loránd Tudományegyetem
Nyelvtudományi Doktori Iskola
suranyianni@gmail.com

Preserving the past, shaping the future: the sixth international workshop on the history of speech communication research (HSCR 2024)

The Sixth International Workshop on the History of Speech Communication Research (HSCR 2024) was held on September 20–21, 2024, at Eötvös Loránd University (ELTE) in Budapest, Hungary. This international workshop focused on advancing speech communication research and provided a unique forum for researchers to integrate past research into future studies.

Jürgen Trouvain, the keynote speaker on the first day, focused on the diversity within phonetic sciences from 1920 to 2020, highlighting the contributions of three influential scholars: Louise Kaiser, Ilse Lehiste, and Anne Cutler. Additionally, he pointed out a gradual increase in female scholars' participation over time.

Michael Ashby discussed the career of Denes (1920-1996), a Hungarian-born engineer who made significant contributions to speech research. Denes' life and work reflect many key developments in speech analysis throughout the second half of the twentieth century. Moving away from analog systems has become standard practice in the field due to his contributions.

Angelika Braun explored the evolution of phonetic research methods and tools in Germany, with a particular focus on the period before World War II. Furthermore, she discussed the contributions of German laboratories to experimental phonetics, highlighting their role in organizing early phonetics conferences and contributing to key journals, such as *VOX* and *Phonetica*.

Annie Rialland explored the early development of experimental phonetics in France, highlighting Pierre-Jean Rousselot's contributions and his collaborations with Adrien Millet and Hubert Pernot, who helped sustain his legacy after he died in 1924. She also explored the historical correspondence between Pernot, Millet, and Josef Chlumský, revealing their efforts to preserve Rousselot's legacy.

Pavel Šturm examined the development of experimental phonetics journals in Europe during the interwar period, relying on the archival correspondence and differing visions of key figures: Hubert Pernot, Edward Wheeler Scripture, Adrien Millet, and Josef Chlumský.

Rüdiger Hoffmann explored the application of photography and cinematography at the Hamburg Phonetics Laboratory, established in 1910. The laboratory has been known for its innovative use of visual media to study speech production. Hoffmann mentioned the photographic archive from the laboratory, now preserved at TU Dresden, which includes over 2,000 digitized images documenting early phonetic experiments, devices, and anatomical studies.

Attila Starčević presented phonetic observations made by György Komáromi Csipkés, a Hungarian theologian, in his 1664 grammar titled *Anglicvm Spicilegium*, which was intended for Hungarian students to learn English and read English texts. Starčević emphasized that although Komáromi was not a linguist, he drew comparisons between English and Hungarian sounds, contributing to our current understanding of some debated issues, particularly regarding the pronunciation of Middle English vowels and consonants.

Quintino Lopes presented the PHONLAB project, which aims to preserve the history of the Experimental Phonetics Laboratory at the University of Coimbra. The project seeks to recover this heritage through the curation of historical instruments, digitization of recordings, and cataloging of documents. Lopes stated that one of the project's goals is to establish a permanent exhibition dedicated to Armando de Lacerda and his work at the University of Coimbra.

Mária Gósy presented the origins, history, and use of the Hungarian word *izé*, as well as its shifting connotations from neutral to vulgar over time, drawing on historical sources, dictionaries, and recordings from the Hegedűs Archive.

András Cser shared his insights into how vowel length was documented and interpreted in Hungarian, particularly from the 16th to 18th centuries, as well as the challenges of accurately representing vowel length in writing, technical constraints, and printing limitations.

On the second day of the workshop, speakers discussed the progression of speech research methods and how these innovations shaped the field. During the first presentation of the day, Ákos Gocsál analyzed the phonetic and linguistic characteristics of early Hungarian broadcast speech. Furthermore, Gocsál argued that creating a database for old newsreels would significantly contribute to a better understanding of historical language use, making them a valuable resource for studying authentic language use that can address specific research questions.

Angélique Amelot presented the process of restoring instruments used in experimental phonetics at the Institute of Phonetics in Paris during the early 20th century. She described the efforts to clean, restore, and identify the instruments and the use of modern techniques, such as 3D printing, to replace missing parts, as well as cataloging and digitally archiving the instruments, making them accessible for the future. Amelot introduced a table game called “Phonetic Pursuit,” a game that displays instruments interactively, allowing participants to engage with the devices and understand their historical use.

Harald Höge presented the SPICOS project, which was the first continuous speech-driven dialogue system developed between 1984 and 1990 by Siemens, Philips, and the IPO Institute of the Technical University of Eindhoven. The SPICOS project laid the foundation for modern AI systems based on data-driven learning and

large amounts of speech resources. Höge also mentioned that the project further influenced commercial applications such as automatic speech recognition systems for mobile phones and medical dictation.

Didier Demolin focused on Pierre-Jean Rousselot's contributions to the evolution of speech sounds and experimental methods in phonetics. Rousselot theorized that phonetic evolution is gradual and that variations in articulation are passed down and modified over generations. Demolin argued that although this theory was not widely accepted during his time, particularly among structuralists, his contributions laid the foundation for the study of sound change and variation, making him a central figure in the development of modern phonetics.

Alexandra Markó provided an overview of the development of articulatory phonetics research in Hungary from the late 19th century to the present. Furthermore, Markó highlighted the most recent phase, beginning in 2016, with the establishment of the MTA–ELTE Lingual Articulation Research Group, which propelled current research using modern tools and methods.

Jingyi Sun presented the academic contributions of Chinese scholars who pursued advanced degrees in phonetics at the University of Paris between the 1920s and the 1940s, as well as the work of these scholars. She further expressed that the work of these scholars continues to inform contemporary research on Chinese tones and that his methodological advancements and empirical findings have had a lasting impact on the field.

Judit Bóna provided a historical overview of child language research presented at the International Congress of Phonetic Sciences (ICPhS) from 1932 to 2023. The analysis shows a steady growth in child language research at the ICPhS, peaking in 2023, although the overall percentage remains low. Bóna emphasized the need for further expansion of child-language research at the ICPhS, and more attention should be given to lesser-studied languages and non-Indo-European linguistic families, as well as to various age groups, particularly infants and adolescents, which are still underrepresented.

Gábor Olaszy presented the development and history of the GOH hearing screening method, which is a tool used to assess hearing in children through the application of low-redundancy synthesized words. Olaszy emphasized that the GOH method is a valuable tool for screening children's hearing, facilitating the early detection of hearing impairments. Its practical use over the past 40 years has helped many children to receive timely treatment. Furthermore, it is an accessible tool that various professionals, such as teachers and speech therapists, can administer. Olaszy also demonstrated the devices to the audience, allowing them to try them out and provide further explanations.

Jacek Kudera presented early studies of South Slavic accentuation, focusing on historical methods of annotating stress through the use of musical notation. He shared his insights into the work of 19th-century linguists and musicologists who attempted to describe the complex tonal features of South-West Slavic languages using musical symbols. Kudera shared his research on how these methods were tested in an experiment involving contemporary native speakers from different dialects (Kajkavian, Čakavian, and Štokavian).

Boróka Balázs presented Gabriella Viktor's pioneering contributions to early Hungarian child language research. Published in 1917, her work *The Language of the Child* is among the earliest scientific explorations of child language in Hungary and is noteworthy for its primarily linguistic focus, diverging from the more common pedagogical and psychological approaches of the time. Balázs concluded her presentation by mentioning that Viktor's work reflected contemporary theories, but also introduced novel ideas, such as the influence of child language on adult speech, a perspective that was later contested.

The Sixth International Workshop on the History of Speech Communication Research (HSCR 2024) brought together researchers from different disciplines to reflect on the evolution of speech communication research. The workshop emphasized the contributions of early pioneers, advancements in methods, and the role of historical preservation. The workshop emphasized the importance of diversity and interdisciplinary collaboration in phonetic sciences, highlighting the need for a broader understanding of time, culture, and methodologies.

NILAY NUR TAŞDEMİR
ELTE BTK Doctoral School of Linguistics
nilaynur@student.elte.hu
<https://orcid.org/0009-0008-8822-1217>

Rebecca Tipton: The Routledge Guide to Teaching Ethics in Translation and Interpreting Education

(London/New York: Routledge. 2024. 236 p. ISBN 9780367565817 [hbk], ISBN 9780367565770 [pbk], ISBN 9781003098485 [ebk])

According to the typology of Saldanha and O'Brien (2014), research in translation and interpreting (TI) can be categorized into four groups: product-, process-, participant-, and context-oriented. Research on ethics in the field of TI falls under participant-oriented research, as pointed out by Chesterman (2018). Translators and interpreters are often confronted during various stages of TI with ethical issues such as the acceptance of a commission or the addition of translators' notes to the text.

Rebecca Tipton asserts that attention should be drawn to ethics explicitly among TI students within tertiary education in order to prepare them for the ethical issues they may face. Such issues may be the impact of subjective personal views or data security, for instance. Her book, *The Routledge Guide to Teaching Ethics in Translation and Interpreting Education*, was published in 2024 as part of the *Routledge Guides in Teaching Translation and Interpreting* series. The main question of the book is how teaching ethics can be effectively integrated into TI training.

To begin with, the book consists of an *Introduction* and five chapters, each with its own subchapters. It can be concluded from the chapters and their topics that Tipton attempts to cover the main pillars of TI education, such as translation, machine translation, interpreting, and research, focusing on TI ethics. This is one of the book's strengths, in addition to its well-founded methodological ground.

In her book, Tipton considers undergraduates, graduates, and doctoral students. Thus, the book addresses the development of ethics in the fields of translation, interpreting, and research. In the brief introduction, Tipton presents the book's aim, which is to assist tertiary-level instructors in integrating ethics into their teaching of TI education. It is emphasized here that a growing interest can be observed in incorporating ethics into TI education; however, there is only a limited number of studies on teaching ethics in TI education.

Chapter 1, *Ethics teaching and teaching ethics*, introduces the basic theoretical background that is necessary in order to include ethics in TI. The writer accentuates that teaching the theory of ethics may be essential, but instructors need to plan their lessons to be task-based and student-centred, defining learning goals accurately. In other words, activities should always be methodologically valid. This chapter is perhaps the most theoretical among the chapters because it introduces the main terms: ethical relativism, ethical competence, ethical sensitivity, ethical knowledge, and ethical judgement.

Chapter 2, *Ethics and the translation curriculum (I)* begins by dealing with the concepts of *integrity* and *thoroughness* for which TI education of ethics should be

responsible. Tipton highlights that a competence-based translation curriculum enhances integrity and thoroughness by forming regular routines in the field of research and revision. Routines help students learn how to organize their work processes. The author includes an exercise adapted from Chesterman's (1997) work for peer translation as an illustration. Students need to write a short essay about a given topic in their mother tongue. They then need to translate each other's text. Afterwards, they provide feedback to each other in pairs. The task is suitable for introducing ethical responsibility, activating moral imagination, and experiencing the consequences of a translator's decisions and risk-taking.

A separate section within the chapter is dedicated to the development of ethical competence with the translation of literary texts. It becomes explicit that the translator's responsibility is extensive, and their decisions during translation may have a profound impact on the perception of other cultures. For instance, inappropriate translation of children's stories can enhance stereotypes and prejudices towards the source language's culture. It is worth noting that one section from the book edited by Pokorn and Koskinen (2020) is recommended as complementary reading in the case study presented in the chapter. This recommendation and the adaptation of Chesterman's (1997) exercise strengthen the link between Tipton's present research and the previous studies in the field of ethics in TI. It is evident that Tipton attentively refers to the preceding theories and research results in the research field of TI ethics.

In the chapter, various topics that educators should address during learning sessions are suggested. These topics can include translator competence, quality, the relationship between the translator and the author, the publisher, the editor, and the reader. Additional topics include class, gender, race, voice, and ethnicity, as well as codes of ethics and theoretical perspectives.

Chapter 3, *Ethics and the translation curriculum (II)*, covers the topic of technology, including issues related to assistive technologies, online environments, ideologies, computer-assisted translation solutions, and collaborative translation. Furthermore, machine translation (MT) is discussed from an ethical perspective. Tipton calls the reader's attention to the risks of applying technology. Educators should explain to students that MT should be used within education and work to an appropriate extent. Understanding should be promoted in connection with the management of data privacy, security, honesty, and copyright.

Tipton's decision to include post-editing in the chapter is understandable, as translators are required to edit and revise texts translated by a neural machine translation service more frequently. Thus, there is a need to include post-editing in TI training. In the future, the chapter could be complemented with a more extensive unit on AI and the ethical issues that may arise from students' use of AI. Moreover, the unethical use of AI within TI training is a relatively new field that is worth mentioning, as institutions may face this problem.

Chapter 4, *Ethics and the Interpreting Curriculum*, is dedicated to the field of interpreting specifically. The chapter examines the connection between ethics and conference and dialogue interpreting, considering them separately. Interpreters often face a significant number of ethical issues during their work. In particular, they need to make decisions in connection with lexis, such as omitting less important information to facilitate better understanding. Tipton describes four situations [court, healthcare, business, media, and social work] in which interpreters need to make ethical decisions in order to thrive toward effective communication. Even if one of the participants fails to realise some vital aspects or details noticed by the interpreter, for example, suppose the interpreter discovers that the woman who has given birth is unaware of the severity of her postpartum condition because the midwife has not recognized the woman's lack of knowledge. The interpreter then intervenes, asking the midwife to include additional information about the condition. Formal education needs to prepare students to make ethical decisions linked to filtering and interventions as well.

Chapter 5, *Teaching research ethics*, ultimately prioritises the procedural and situational practices within the teaching of research ethics. In the chapter, *inter alia*, the topics of integrity and risk during the research process, as well as data collection and storage, are discussed from an ethical perspective. A brief section can be read as an introduction to research ethics in general, followed by an examination of the challenges of teaching ethics.

A separate chapter describes the central notions with reference to research ethics and their impact on TI researchers. While students at lower levels of tertiary education can conduct various types of research, the entire chapter may be more relevant to educators who teach doctoral students or are involved in PhD programmes as supervisors. Students need to be aware that it is crucial not to cause harm to research participants, to obtain informed consent from them, and to prioritise confidentiality, data management, storage, and retention.

To conclude, it can be stated that Tipton's book is a comprehensive work on teaching ethics in TI education and research. The chapters are mindfully structured and compiled. The chapters provide guidelines for educators, including theoretical parts as well. Some TI institutes or departments may intend to include TI ethics in their training as separate courses. It would be useful if sample syllabi for at least one lecture and one seminar were added to the second chapter.

Educators may be able to include and teach ethics in TI education sessions with the help of this volume. The book can be recommended to incorporate ethics into courses, lessons, curricula, or training. Ethical issues related to the use of recommended activities, prompts, and questions can be discussed within educational sessions.

References

- Chesterman, A. (1997). *Memes of translation: The spread of ideas in translation history*. John Benjamins. doi: <https://doi.org/10.1075/btl.22>
- Chesterman, A. (2018). Translation Ethics. In L. D'hulst and Y. Gambier (Eds.), *A History of Modern Translation Knowledge. Sources, Concepts, Effects* (pp. 443–448). John Benjamins. doi: <https://doi.org/10.1075/btl.142.62che>
- Pokorn, N. K., & Koskinen, K. (2020). The ethics of linguistic hospitality and untranslatability in Derrida and Ricœur. In K. Koskinen and N. K. Pokorn (Eds.), *The Routledge handbook of translation and ethics* (pp. 87–98). Routledge. doi: <https://doi.org/10.4324/9781003127970>
- Saldanha, G., & O'Brien, S. (2014). *Research methodologies in translation studies*. Routledge. doi: <https://doi.org/10.4324/9781315760100>

ADRIENN BIRÓ

Eötvös Loránd University
Department of Translation and Interpreting
Translation Studies Doctoral Programme
adriennbiro@student.elte.hu
badrienn1996@gmail.com
<https://orcid.org/0009-0001-7896-2387>