

Baditzné Pálvölgyi Kata¹

Eötvös Loránd Tudományegyetem
Bölcsészettudományi Kar
Szakmódszertani Központ – Romanisztikai Intézet

Mesterséges intelligenciával létrehozott hangzóanyagok vizsgálata a magyar orvosi szaknyelv oktatásában

<https://doi.org/10.48040/PL.2026.1.8>

Kutatásunk azzal a céllal született, hogy a magyart mint idegen nyelvet oktatók számára rövid áttekintést nyújtson azokról az akár ingyenesen is hozzáférhető mesterséges intelligencián (MI) alapuló lehetőségekről, amelyek hangzóanyagok létrehozása által hatékonyan támogatják és megkönnyítik a nyelvtanár munkáját. Elsősorban az orvosi szaknyelvoktatás szempontjából vizsgáljuk meg a jelenleg elérhető opciókat, azonban a magyart mint idegen nyelvet bármilyen szinten és kontextusban tanító szakemberek számára is releváns megállapításokat teszünk. A kutatás azt vizsgálta, milyen színvonalon teljesít az MI a következő területeken a magyar orvosi szaknyelv vonatkozásában: 1) szakszövegekből az oktató által készített hangzóanyagok, valamint 2) az autonóm tanulásra szolgáló, speciálisan orvosi kontextusra optimalizált, hanganyagokat önállóan létrehozó vagy akár hangüzenetek váltására is képes chatbotok. Az eredmények alapján a hanggenerálás (beleértve a hangklónozással létrehozott mintákat is) még fejlesztendő terület, az alkalmazások a természetes beszéd hangleadását kevésbé képesek előállítani, de összességében a szegmentális és szupraszegmentális jellemzők vizsgálatakor az ElevenLabs emelkedett ki a többi hasonló alkalmazás közül. Az autonóm nyelvtanulásra szolgáló chatbotok is csak körültekintéssel alkalmazhatók egyelőre, a kiejtés tekintetében még hiányosságok tapasztalhatók, az alkalmazott szókincsre viszont inkább betegközpontú, laikus regiszter a jellemző.

Kulcsszavak: magyar orvosi szaknyelv, szaknyelvoktatás, mesterséges beszéd, autonóm tanulás, chatbotok

Bevezetés

A mesterséges intelligencia (MI) szerepe a felsőoktatásban egyre jelentősebb, különösen a személyre szabott oktatás lehetősége és az adminisztratív folyamatok hatékonyabbá tétele révén (Robert, 2024; Harris, 2024). Az etikai irányelvek betartásának nehézségei, a visszaélések lehetősége és a digitális szakadék problematikája ugyanakkor mindenképpen nehezítő körülménynek számítanak (Freeman, 2024; Schroeder, 2024). Az MI sikeres integrálása érdekében elengedhetetlen mind az oktatók, mind a hallgatók felkészítése és a tudatos implementáció. A hallgatók nyitottak az MI-eszközök szélesebb körű alkalmazására, különösen a nyelvtanulás, szövegfordítás, stilisztikai korrektúra és az anyaggyűjtés terén (Rajki et al., 2024). A felhasználók szempontjából viszont egyértelműen kihívást jelent a hirtelen megjelent alkalmazásdömping (Horváth, 2023), nehéz kiválasztani az igényeinkhez leginkább megfelelő platformokat. A magyar munkanyelvet használók ma még inkább elutasíthatják az MI mindennapi alkalmazását a kipróbálást követő kezdeti csalódások miatt, mivel az általa generált magyar nyelvű kimeneti tartalom minősége gyakran elmarad a várakozásoktól, különösen például az angol nyelvű outputokkal összehasonlítva.

A tanulmány célja, hogy a magyart mint idegen (szak)nyelvet oktatók számára rövid áttekintést nyújtson azokról a lehetőségekről, amelyek hatékonyan támogatják és megkönnyítik a nyelvtanár munkáját a mesterséges intelligencia segítségével generált hanganyagokkal kapcsolatban. Elsősorban az orvosi szaknyelvoktatás szemszögéből elemezzük a rendelkezésre álló felületeket, a kutatás azonban a magyart mint idegen nyelvet

¹ Baditzné Pálvölgyi Kata: <https://orcid.org/0000-0002-7106-7427>

bármilyen szinten és kontextusban tanító szakemberek számára releváns tanulságokkal szolgálhat.

Bár számos tanulmány foglalkozik a mesterséges intelligencia nyelvoktatásban betöltött szerepének terjedésével (Kohnke et al., 2023), nem áll rendelkezésre a magyar mint idegen nyelv tanítását támogató és sikerrel alkalmazható megoldásokról átfogó munka, pedig nagy szükség volna rá. Az MI már ma is hatékonyan támogathatja a magyar mint idegen nyelv tanár munkáját a szakmai szövegek előállításában, fordításában és lektorálásában, a magyar nyelvű audiovizuális és hangzóanyagok előállításában és az ezekhez szükséges átiratok készítésében, prezentációkészítésben és az autonóm tanulást segítő chatbot típusú alkalmazásokban is.

Ebben a tanulmányban arra a kutatási kérdésre keressük a választ, hogy a magyar orvosi szaknyelvoktatás támogatására mennyire megfelelőek jelenleg a mesterséges intelligenciával generált hangminták. A következő két területre fókuszálunk: (1) magyar nyelvű hangzóanyagok előállítása konkrét szaknyelvi input alapján és (2) az autonóm szaknyelvtanulást segítő felületek, chatbotok alkalmazása. Természetesen a kutatás nem lehet teljes körű, már a tanulmány terjedelmi korlátai miatt sem, célunk, hogy ezeken a területeken betekintést nyújtsunk a lehetőségekbe és megvizsgáljunk a jelenleg könnyen hozzáférhető felületeket, illetve felhívjuk a figyelmet azokra a szempontokra, amelyek alapján érdemes a jövőben is ezeket az alkalmazásokat saját használatra tesztelni. Minden esetben csak olyan alkalmazásokat tekintünk át, melyeknek van legalább „freemium” verziója, tehát, ha korlátozottan is, de a vizsgált funkciók ingyenesen is rendelkezésre állnak. Kizárólag az asztali gépen is elérhető alkalmazásokkal foglalkozunk, mert a mobil eszközökre optimalizált változatok esetenként kevésbé kényelmesek – például tananyagfejlesztés szempontjából.

A mesterséges hangok elvárt jellemzői

A mesterséges intelligencia segítségével generált hangok ma már annyira hasonlíthatnak az emberi hangokra, hogy szinte lehetetlen felismerni, emberi beszédet hallunk-e, vagy gép által szintetizáltat (Barrington et al., 2025). Az oktatóanyagok esetében hatalmas előnyt jelenthet, hogy jóval olcsóbbak a gépi hangok, mintha professzionális emberi hangon mondanák fel a szöveget és nincs szükség időpontegyeztetésre a színészekkel, de az egyértelműen mesterséges, nem természetes emberi beszédnek ható hangminták elutasítást válthatnak ki a hallgatóból, ahogy ez például a reklámparban is használt mesterséges hangok tesztelése kapcsán kiderült (NielsenIQ, 2024). Épp ezért érdemes megvizsgálunk, mitől lesz természetes egy Text-to-Speech (TTS), tehát gép által generált beszédminta.

Egy adott nyelven természetesnek ható beszédminta számos, alapvetően nyelvi kritériumnak kell, hogy megfeleljen (Wang–Székely, 2024; Chan–Kuang, 2025; Li et al., 2023), tekintsünk most át néhányat. Elsőként említhetjük a szegmentális jegyeket, mint például a hangok tiszta artikulációja, a nyelvre jellemző magánhangzó-színezet megtartása és a koartikuláció jelenléte. Ezután következhetnek a szupraszegmentális jegyek, mint például a hangsúlyjellemzők, vagy az egyes nyelvekben jellemző átszótagolódás: átlagos tempójú beszéd közben is gyakori jelenség, hogy a szóvégi mássalhangzót átszótagolják a következő, magánhangzóval kezdődő szóhoz, mint például a francia *avec elle* ‘vele (nőnemű)’ [avɛk ɛl] → [a.vɛ.kɛl] példában látható. Ezen kívül fontos szupraszegmentális szempont a nyelv jellemző ritmusának megtartása (különböző nyelvek átlagos beszédtempója nem egyforma, Goldman-Eisler, 1968; idézi Gósy, 2004:205), valamint a nyelvspecifikus intonációs mintázatok (2024-ben több vizsgált bot még nem volt alkalmas konzisztensen a magyar emelkedő-eső dallamív létrehozására a semleges eldöntendő kérdésekben, vö. Baditzné Pálvölgyi, 2024). Nem feledkezhetünk el azokról a megakadás-jelenségekről sem, amelyek egy adott nyelv sajátjai, mint például a hangmegnyúlások vagy a kitöltött szünetek: nem minden nyelvre jellemző ugyanazoknak a hangoknak a megnyújtása – míg a szóvégi *-l* nyújtása természetes a

spanyolban, a magyarban kevésbé (Baditzné Pálvölgyi, 2020) –, és a kitöltött szünetek sem történnek ugyanannak a magán- vagy mássalhangzónak a megnyújtásával (magánhangzók tekintetében a magyarra az „őzés” jellemző, míg az angol nyelv a svát [ə] használja kitöltő hangként). Ugyanakkor a TTS hangnak nem elég megfelelő szegmentális és szupraszegmentális jegyekkel és hezitációs mintázatokkal rendelkeznie, adott esetben érzelmek kifejezésére is alkalmasnak kell lennie, robotikus rezonanciák, „fémcsengés” nélkül, illetve, mivel maguk az emberi hangminták sem tökéletesek, kismértékű tökéletlenség látszatát kell kelteniük.

Adott szövegből mesterséges hangot előállító alkalmazások tesztelése

Saját hanganyagok előállítása általában akkor merül fel, ha nem áll rendelkezésre az általunk használt tankönyvben elégséges hangfelvétel, ami szaknyelvoktatás esetén különösen gyakori helyzet, hiszen jóval kevesebb szaknyelvkönyv van a piacon, mint általános nyelvkönyv. A szaknyelvoktatók jellemzően nagyon specifikus szövegekkel dolgoznak, amelyekhez nem könnyű létező, autentikus hanganyagot is találni, így mindenképp nagy előny, ha az oktatók személyre és saját csoportjuk konkrét igényeire szabott szövegekből készíthetnek hanganyagot.

Felmerülhet a kérdés, miért lehet szükség magyar mint idegen (szak)nyelvórán mesterséges hangok létrehozására és felhasználására, a magyar anyanyelvű tanárok ugyanis könnyen előállíthatnak maguk is felolvasással hangzóanyagokat, nem lenne szükséges mesterséges intelligenciával előállítani ezeket. Az egyik ok lehet az, hogy ma már számos környezetben (pl. egy CT működése közben is) találkozhatunk mesterséges, gépi hangokkal, tehát a nyelvtanulónak a gépi hangot is értenie kell célnyelven. Amennyiben teljesen emberi hangmintákat szeretnénk generáltatni, annak is lehet azonban több oka. Nem szerencsés például, ha csak a nyelvtanár kiejtését ismerik a tanulók, és mesterséges hangokkal viszont többféle kiejtéssel mondatható fel a célnyelvi szöveg. Ezek a hangminták alapértelmezetten mentesek a háttérzajtól, a túlzott mennyiségű megakadásjelenségtől, hezitálástól, így nem kell újra rögzíteni vagy megvágni a felvételt, ha a felolvasó megakad. Ezek a mesterséges hangminták többnyire ugyanakkor érzelmektől is mentesek, így a legtöbb elérhető alkalmazás leginkább tényszerű monológok felmondására használható.

Az alábbiakban azt a vizsgálatot mutatjuk be, amelyben azt teszteljük, hogy négy, adott input alapján mesterséges hang generálására is alkalmas felület (az MS Word, a ChatGPT, az Elevenlabs és a Suno) milyen mértékben képes magyar szövegeket megfelelő szegmentális és szupraszegmentális jegyekkel előállítani, képes-e természetes módon hezitálni, illetve azt is vizsgáltuk, hogy miként kezeli a magyar szövegbe ékelt idegen szavakat. Az alkalmazásokat előzetes nyelvtanári tapasztalataink és pedagógusok preferenciái (Karageorgakis, 2023) alapján választottuk ki. Az MS Word és a ChatGPT esetében felolvastathatunk az alkalmazásokkal, de a hangok rögzítését saját magunknak kell megoldanunk. Képernyőfelvételt készítünk az alkalmazások által felolvasott szövegről, és ezt az MP4 formátumban mentett fájlt konvertáljuk MP3-má (Baditzné Pálvölgyi, 2024). Az Elevenlabs felületén azonnal MP3 formátumban tölthetjük le a generált hangmintát. A Suno pedig nem egyszerűen hangmintát, hanem egy tetszőleges stílusban megzenésített MP3 fájlt hoz létre. Így ugyan nem életszerű szakmai anyagokat hozhatunk létre, de a szakszövegek megzenésített változatával színesíthetjük a tanórákat.

Ebben a 2025-ben végzett kutatásban a MS Word alkalmazás alapértelmezett férfi hangját választottuk, a ChatGPT 40 mini verzióban elérhető „Spruce” hangot. Az Elevenlabs alkalmazásban elérhető „Magyar férfi – Hungarian Male, Young Hungarian Storyteller” hangot (a programon belüli azonosító kódja *TumdjBNWanlT3ysvcIWh*) használtuk. A Suno felületén csak a stílust választottuk meg (cumbia műfajt), a generált mintákban (az alkalmazás alapértelmezetten két rövid zeneszámmal készít) férfihangot kaptunk, így a beszélő neme mind a

négy alkalmazásban megegyezett. Minden esetben az 1. táblázatban feltüntetett mondatokat teszteltük:

1. táblázat. Az alkalmazásokban tesztelt szaknyelvi szöveg átirata

<i>Pajzsmirigy-alulműködése van? Feltehetőleg dzsungelláza van. Tele van a polc madzagokkal. Ez egy Riester márkájú otoszkóp. Nyújtsa ki és hajlítsa be a karját! A bátyámnak zsírmája volt.</i>	<i>A beteg refluxos, és krónikus hmmm bronchitise van. Mmmm, sajnos a tüdő lyukas. Küldje be a következőt! Alacsony a vérnyomása. Szedjen C-vitamint!</i>
---	---

A felolvasztatott mondatok összeállításakor arra törekedtünk, hogy a következő jelenségek mindegyike szerepeljen bennük: a szegmentális jegyek tekintetében a magánhangzók és mássalhangzók ejtése (így a magyar hangzókészlet minden eleme szerepelt), a szegmentális jegyeken túl hangkapcsolatok ejtése (hasonulások), illetve kijelentő, felszólító mondatok és eldöntendő kérdések is (a hanglejtés ellenőrzése miatt). Mivel a hezitálás tekintetében korábbi tapasztalatok alapján (Baditzné Pálvölgyi, 2024) az MI-alapú hanggeneráló alkalmazások nem teljesítettek jól, illetve az adott hangmintán belüli nyelvváltások esetében sem kaptunk megfelelő eredményeket, ebben a kutatásban is vizsgáltuk a célnyelvi mondatokba beékelte időnyerő megakadásjelenségeket (ezeket „hmmm” vagy „mmmm” betűkkel írtuk át), valamint az idegen szavak megvalósulását (ez utóbbiakra főleg a szaknyelvben előforduló idegen szavak magas aránya miatt volt szükség).

A 2. táblázat tartalmazza a négy alkalmazás tesztelésekor kapott eredményeket. Az első oszlopban a tesztelt jelenségek láthatók, a másodikban pedig a szám, ahány esetben a vizsgált jelenség előfordult összesen. A következő négy oszlopban a számadatok azt mutatják, milyen százalékban kaptunk elfogadható megoldást a hangmintában. A magánhangzók esetén a *Riester* szóban az *ie* betűk egy magánhangzót, az [i]-t jelölték, a mássalhangzók esetén a hasonulások következtében keletkezett hangokat vettük figyelembe (tehát a *nyújtsa* szóban a *ts* egy [tʃ] hangnak felelt meg, nem [t] + [ʃ] mássalhangzóknak), és a hezitálások kitöltött szüneteinek mássalhangzóit nem vettük figyelembe a számításokor, ezeket külön kategóriában vizsgáltuk. Végül megvizsgáltuk az egyes hangminták beszédtempóját, az első kiejtett hangtól az utolsóig mérve szünetekkel (szótag/másodpercben mérve), hogy összevessük a szakirodalom alapján a magyarban átlagosnak tartott 5,97 szótag/másodperc értékkel (Gósy, 1988; idézi Kassai, 1993:69). Egyik alkalmazásban sem változtattunk az alapértelmezetten beállított beszédtempón a tesztelés alkalmával (habár ez az MS Wordben és az Elevenlabsben is elérhető opció, mely kezdő csoportoknál kimondottan hasznos funkció lehet).

2. táblázat. A hangmintákban tesztelt jelenségek megvalósítása négy hanggeneráló MI alkalmazásban

Tesztelt jelenségek	előfordulás (db)	sikeres megvalósítás %-ban			
		MS Word	ChatGPT	Elevenlabs	Suno
magánhangzók kiejtése	102	100	98	100	86
mássalhangzók kiejtése	141	99	96	99	96
hangkapcsolatok ejtése (hasonulások)	4	100	100	100	75
szóeleji hangsúly	50	100	100	100	98
kijelentő mondatok hanglejtése	7	100	100	100	100
felszólító mondatok hanglejtése	3	100	100	100	100
eldöntendő kérdések hanglejtése	1	0	0	0	0
hezitálás	2	0	100	100	100
beékelte idegen szavak kiejtése	1	0	100	100	100
beszédtempó (szótag/mp; magyar átlag 5,97)		4,01	5,7	4,07	2,92

Az összefoglalóból az derül ki, hogy a szegmentális jegyek tekintetében nem minden alkalmazás teljesített hiba nélkül. A ChatGPT modelljében még folyamatosan érzékelhető volt az angol kiejtés hatása elsősorban a mássalhangzók esetében, például szóeleji hehezetes ejtés: *tele* [t^hɛlɛ], az r-ek nem pergők, hanem lebegők, mint az amerikai angolban (*refluxos* [ˈɹɛfluksɔʃ]), illetve a magánhangzók színezete esetenként diftongizálódást mutatott, például *krónikus* [ˈkɹaoːnikʊʃ], *vérnyomása* [ˈweɪnomaːʃɒ]. A *dzsungelláz* szót pedig, valószínűleg ritka előfordulása miatt, helytelenül [dʒ̃]-vel ejtette a program. A Suno alkalmazás kapcsán általános tapasztalatként elmondható, hogy nem minden esetben ismeri fel azonnal, milyen nyelven kellene az outputnak megjelennie. Esetünkben a *pajzsmirigy* szó után már majdnem teljesen a magyar hangzókészlet alapján ejtette a szavakat, de előfordult az *a* [ɒ] és az *e* [ɛ] hangok helyett inkább spanyol [a] és [e] ejtése, a *dzs* [dʒ] helyett [dʒ̃], *gy* [j̃], *ly* [j] helyett [gi], [li] ejtés valósult meg, és az *otoszkóp* szó esetében szóvégi hangsúly volt hallható, azonban ez a műfaji sajátosságoknak is betudható (ebben a latin zenei műfajban jellemző a sor végén a hangok elnyújtása, ami okozhat hangsúlypercepciót). Mind a négy vizsgált alkalmazásban [ʃ]-sel történt a szóvégi [s] ejtése a *bronchitis* szóban.

Egyik vizsgált alkalmazásban sem volt megfelelő az eldöntendő kérdések hanglejtése: az utolsó előtti szótagig emelkedő, majd onnan eső dallamot (Varga, 1994) nem sikerült reprodukálni, az utolsó előtti szótagon nem jelentkezett emelkedés, hanem esés. Hezitálás tekintetében az MS Word sajnos a hezitálásra alkalmazott betűket olvasta össze (de betűszóként, tehát pl. [ememem]), a beékelt idegen szavakat pedig részben magyar kiejtéssel reprodukálta. A beszédtempó esetén egyik alkalmazás sem érte el az átlag értéket, de a legközelebb hozzá a ChatGPT alkalmazás hangmintája állt. A Suno kevesebb, mint a fele értékkel a legalacsonyabb, azonban ez kevésbé objektív adat önmagában, hiszen ott a zenekíséret miatt hosszabb szünetek voltak a felvételen. Látható tehát, hogy az Elevenlabs a legjobban teljesítő alkalmazás a vizsgáltak közül, csak a már említett emelkedő-eső hanglejtés esetében tapasztalhattunk anomáliát.

Mesterséges hangok minősége autonóm tanulást támogató felületeken

Az autonóm tanulás a tanuló önrányított és önálló tanulását jelenti, ahol a tanuló aktív résztvevője saját fejlődésének (Holec, 1981; Little, 1991). Esetünkben olyan MI-alapú megoldásokat vizsgálunk, amelyek használatával akár oktatói közreműködés nélkül egy orvostanhallgató magyar orvosi szaknyelvhasználatot igénylő szituációban gyakorolhat hangalapú csevegésben az MI-vel, vagy önfejlesztésre használhatja az MI által generált hangzóanyagot.

Az orvosképzésben vannak már kezdeti pozitív tapasztalatok az olyan MI chatbotok alkalmazására, amelyek támogathatják az orvos–beteg kommunikáció autonóm fejlesztését. Chiu és munkatársai (2025) a rossz hír empatisz közlésének gyakoroltatásához használták a ChatGPT-t sikerrel texasi orvostanhallgatók képzésében, Amithasagaran és munkatársai (2025) vizsgálatai alapján pedig szintén texasi orvostanhallgatók ígéretesnek tartották és biztonságos kiegészítő eszköznek tekintették az MI- és virtuális valóságalapú chatbotot orvos–beteg interakciók szimulálására.

A magyar orvosi szaknyelv vonatkozásában még várnunk kell hasonló kutatásokra, ezért a következő alfejezetben azt vizsgáljuk meg, milyen színvonalon teljesítenek orvosi szaknyelvi környezetben a magyar nyelven is elérhető, autonóm tanulást lehetővé tevő chatbotok, valamint egy podcastgeneráló felület, a Notebook LM. Előző vizsgálatunkhoz képest az a különbség, hogy ebben az esetben csak a témát, kontextust szolgáltatja az oktató, de nincs konkrét input, ami a hangzóanyagban szóról szóra felolvasásra kerülne.

Először három, önálló nyelvgyakorlást is lehetővé tevő chatbotot vizsgáltunk meg. Az elemzett alkalmazások a Mizou, a Gliglish és a ChatGPT hangalapú csevegő funkciója voltak,

az alkalmazásokat Russel Stannard (2024) MI-szakértő nyelvoktató módszertani ajánlásai és saját pozitív oktatói tapasztalataink alapján választottuk. A Mizou esetén olyan saját chatbotot hoztunk létre promptokkal, amelynek működése orvos–beteg párbeszéd szimulációján alapul. A következő, angol nyelvű, maga az alkalmazás által a témához javasolt prompt alapján készült a saját bot:

„You are a medical professional chatbot. The student is a medical student practicing patient communication. Engage in a dialogue about the patient's complaints, focusing on empathy and clarity. The student should ask questions and respond to the patient's concerns, simulating a real consultation. The expected outcome is improved communication skills and empathy in patient interactions”.

A Gliglish felületén eleve választható az orvos–beteg szimulációt gyakorló chatbot, a ChatGPT pedig a következő utasítást kapta: „Most orvosi rendelőben szimulációt szeretnék. Te leszel az orvos.” Mindhárom esetben ugyanazt a bemeneti szituációt kapták a botok: a beteg talpfájdalomra panaszkodott, a párbeszédet pedig mobiltelefonon képernyőfelvétellel rögzítettük. A párbeszédéből a 3. táblázatban részleteket közlünk, amelyekben P. a páciens, C. a chatbot beszédfordulóit jelzi. A párbeszédet átírtait a Turboscribe programmal végeztük, amely a magyar nyelv átírásához is már nagy pontossággal alkalmazható.

3. táblázat. Autonóm tanulást segítő chatbotokkal folytatott fiktív orvos–beteg párbeszéd részletei

Mizou	Gliglish	ChatGPT
C.: (Üdvözöllek! Készen állsz a beteg párbeszéd szimulációjára?) P.: Igen. Én vagyok a beteg, és akkor szeretném azt mondani, hogy nagyon fáj a talpam. C.: Rendben, Kata, téged figyelmesen hallgatlak. Mikor kezdődött a talpfájdalom, és tudnál mesélni arról, hogy mivel állt összefüggésben? Van esetleg olyan tevékenység, ami súlyosbítja a fájdalmadat? P.: Körülbelül egy hónapja kezdődött, és általában este érzem ezt. C.: Köszönöm, hogy megosztottad velem, Kata. Az, hogy este érzed a fájdalmat, sokat elmond. Milyen jellegű a fájdalom? Éles, tompa, vagy esetleg szúró? Esetleg észleltél más tüneteket is, mint a duzzanat, vagy a bőrpír? P.: Ilyen égő fájdalom, és nem éreztem más tüneteket. C.: Az égő fájdalom fontos információ (...).	C.: (Jó reggelt! Dr. Feelgood vagyok. Miben segíthetek ma?) P.: Jó napot! Az a baj, hogy nagyon fáj a talpam. C.: Sajnálom, hogy fáj a talpad. Mikor kezdődött a fájdalom? Milyen tüneteket tapasztalsz? P.: Egy hónapja kezdődött körülbelül, és esténként szokott fájni. Olyan égő fájdalom. C.: Értem. Az égő fájdalom fontos. Van más tüneted is? Például duzzanat vagy bőrpír? P.: Nem, nincs más tünetem. (...). C.: Köszönöm az információt. Kérlek, meséld a kórtörténetedről. Szedsz valamilyen gyógyszert? P.: Csak vérnyomáscsökkentőt.	P.: Azzal a problémával jönnék, hogy fáj nagyon a talpam. C.: Jó napot kívánok! Értem. Pontosán hol fáj a talpa? Az egész talp, vagy inkább csak egy része, például a sarka vagy a lábujjak környéke? P.: Az egész talpam fáj, de főleg csak esténként. C.: Értem, tehát főleg este jelentkezik a fájdalom. A nap folyamán nem érez fájdalmat, csak, amikor pihenni próbál. Van valami, ami enyhíti a fájdalmat, például felpolcolás vagy pihenés? P.: Nem, egyik sem (...). Esetleg rövid ideig, hogyha jég ér hozzá. C.: Értem. A jég enyhíti, de csak rövid ideig. Ez arra utalhat, hogy gyulladás van a talpban.

Mindhárom részlet a chatbot első három megnyilatkozását tartalmazza (a köszönést leszámítva), ezeket fogjuk elemezni a következő szempontokból: hány szegmentális, illetve szupraszegmentális hibát tartalmazott az adott minta? Mekkora a bot megnyilatkozásainak beszédtempója? Mivel a generált tartalmat nem tudtuk előre szabályozni és a három mintában garantálni konkrét vizsgálandó nyelvi jelenségeket (mint például közbeékelt idegen szavakat), ezért további, kiejtéssel kapcsolatos aspektusokat nem vizsgáltunk, azonban olyan további nyelvi jellegzetességeket igen, amelyek relevánsak lehetnek a szaknyelvoktatás szempontjából.

Az orvosi szaknyelv különböző rétegeit legfőképpen a konkrét nyelvhasználat színtere határozza meg: a tudományos szaknyelv a legmagasabb, absztrakt szint; a gyakorlati szakmai nyelvhasználat a gyógyítók egymás közötti szakmai érintkezésére szolgál, míg a szakemberek és laikusok közti kommunikációban közvetítőnyelvről beszélhetünk (Kuna, 2024:20-25). Esetünkben a chatbotoktól az volt elvárható, hogy természetes orvosbeteg kommunikációt szimuláljanak. Ezt három szempontból vizsgáltuk:

- 1) a dialógus formalitása (tegezi, vagy magázza-e a páciens a bot);
- 2) a szaknyelvhasználati réteg (használ-e releváns szakszavakat; ez esetben az anamnézis-felvétellel, diagnózissal, illetve orvosi javaslatokkal kapcsolatos szavakat tekintettük annak), ezek hivatalos szaksargon (tudományos szaknyelv) vagy laikusokkal szemben elvárható köznyelvi „betegbarát” megfelelők-e (közvetítőnyelv), illetve
- 3) találunk-e olyan fatikus diskurzusjelölőket, amelyekkel a beteget az orvos arról biztosítja, hogy figyel rá.

4. táblázat. A chatbotokban vizsgált nyelvi jelenségek

Vizsgált aspektusok		Mizou	Gliglish	ChatGPT
szegmentális hibák száma		0	0	4
szupraszegmentális hibák száma (beleértve: hangkapcsolatok)		4	2	3
beszédtempó (szótag/perc)		5,33	3,29	4,86
tegezés/magázás		tegezés	tegezés	magázás
szakszavak	laikus	talpfájdalom, súlyosbít(ja), fájdalm(adat), éles, tompa, szűrő, tünet(ek), duzzanat, bőrpír	fájdalom, tünet(eket), fájdalom, tünet(ed), duzzanat, bőrpír, gyógyszer(t)	fájdalom, enyhít, főlpolcolás, gyulladás
	szakzsargon	-	kórtörténet(edről)	-
fatikus elemek		rendben, Kata, téged figyelmesen hallgatlak, Köszönöm, hogy megosztottad velem, Kata. ..., sokat elmond. ..., fontos információ.	sajnálom, hogy... értem Köszönöm az információt	értem (3x)

Az adatok alapján kijelenthető, a Mizou és a Gliglish nem vétenek szegmentális kiejtési hibákat, azonban a ChatGPT igen: négy esetben is inkább angol kiejtést tapasztalhattunk, a következő szavakban: *pontosan* (hehezetes [p^h]), *sarka* (hehezetes [j^h]), *környéke* (*lebegő* r [ɹ]), *talpban* (hehezetes [t^h]). Jelenleg mindhárom vizsgált alkalmazás esetében probléma, hogy az eldöntendő kérdések hangleyjtése inkább hasonlít a kijelentő mondatokéhoz, egyetlen esetben sem sikerült az elvárt emelkedő-eső dallamot reprodukálni. A beszédtempó esetén pedig elmondható, hogy lassabbak az átlagosnál (legkevésbé a Mizou).

Ami a tartalmi szempontokat illeti, a párbeszédekből látszik, hogy mind a Mizou, mind a Gliglish alapértelmezetten tegezi a páciens, ami magyarországi kontextusban felnőtt beteggel nem jellemző. A szakszavak a Gliglish által használt *kórtörténet*en kívül laikus regiszterben maradnak, és mindhárom alkalmazás használ a kapcsolat fenntartására szolgáló megerősítő fatikus elemeket, de a ChatGPT nagyon korlátozottan, csak egyetlen ilyen elemet ismételt (értem). Jellemző a kontextustól idegen angol tükörfordítás néhány szerkezetben: *téged figyelmesen hallgatlak*; (*Az, hogy este érzed a fájdalmat,*) *sokat elmond*; (*Az égő fájdalom*) *fontos információ*.

Az autonóm tanulást természetesen nem csak chatbotok segítségével támogathatjuk, a Google 2024 óta már nem tesztüzemmódban futó Notebook LM alkalmazása képes arra, hogy általunk feltöltött szakmai anyagok, például orvosi szakkönyvek alapján kivonatot készítsen, vagy kérdések-válaszok generálásával feldolgozza a tananyagot, akár podcast formájában is.

Az alapértelmezetten angol nyelven elérhető podcast igen hasznos tanulósegéd lehet, de sajnos magyar nyelvű minta 2025-ben még nagyon kezdetleges minőségben és csak az utasításokban explicit módon többször kért „kizárólag magyar nyelven” paranccsal volt generálható. A következő példa egy spanyol nyelvű traumatológia tankönyv (Orrego–Morán, 2014) egyik fejezete alapján készült magyar nyelvű podcast részlete (az átiratot ismét a Turboscribe alkalmazás segítségével készítettük, amelyben a kiejtést hűen tükröző megoldást kapunk). Mivel szaknyelvkönyvről van szó, tudományos szaknyelvvvel fogunk találkozni, a podcastben megjelennek orvosi terminusok (**metakárpus*, *politraumatizált*), azonban ezek nem mind felismerhetők a fonetikai torzulás következtében. A táblázatban az átirat alatt kigyűjtöttük a kiejtési hibákat, külön feltüntetve azokat, amelyek a szövegértést is akadályozzák.

5. táblázat. A Notebook LM által generált podcast részletének átirata és a kiejtési hibák

Az átirat szövege	<i>A kezdeti röntgén felvételen sokszor nem is látszik a törés. A tünetek ellente enyhék lehetnek, könnyen összetévezhetőek simán rándulással. Ezért fontos, hogy komolyan vegyük a csuklósérüléseket, és ha fájdalom nem múlik, keressünk fel szakarvost. Akkor most térjünk át a metakárpus törésekre. Azok a minnyire gyakoriak. A metakárpusok a kész középcsontok, amik az ujják és a csukló között helyezkednek el. A törések a második leggyakoribb készsérülés, és ekkor előfordulnak politraumatizált betegeknél, akiket szeret több sérülést szemben. Rendszerben, akkor most vegyük golcsó alá az interfele-angyalis ízületi sérüléseket.</i>
kiejtési hibák	* <i>röntgén</i> , * <i>tünyetek</i> , * <i>ellente (eleinte)</i> , * <i>összetévezhetőek</i> , * <i>csuklósírüléseket</i> , * <i>szakarvost</i> , * <i>metakárpus (metacarpus)</i> , * <i>a minnyire</i> , * <i>kész (kéz)</i> , * <i>ujják</i> , * <i>golcsó</i>
értelmezavaró kiejtési hibák	* <i>simán (sima)</i> , * <i>ekkorán (?)</i> , * <i>szeret (szerez?)</i> , * <i>szemben (?)</i> , * <i>interfele-angyalis (?)</i>

Látható, hogy a szavak kiejtése gyakran bizonyos szegmentális jellemzőkben eltér a célnyelvitől, ez gyakran a magánhangzó-színezetet érinti csak (pl. *röntgen* ['røntge:n], *ujják* ['ujja:k]), vagy egy-egy mássalhangzót (*tünetek* ['tynetek]), amelyek a szó megértését még nem befolyásolják, de néhány esetben az eltérés a szövegértést is ellehetetleníti (pl. a *szemben* szó a mondatban nem értelmes), és feltételezhetően nem is csak kiejtési, hanem fordítási hiba.

Konklúzió

A tanulmányban két fő oktatási területen tekintettük át a jelenleg akár ingyenesen hozzáférhető mesterséges intelligencián alapuló alkalmazások magyar nyelven is elérhető funkcióit az orvosi szaknyelvoktatás vonatkozásában: hangzóanyagok készítése megadott konkrét szaknyelvi forrásból, és autonóm tanulást támogató megoldások hanggeneráló funkciói adott szaknyelvi téma/kontextus alapján. Célunk nem csupán az volt, hogy feltérképezzük, milyen fejlettségi szinten áll az MI 2025-ben a magyar nyelvű hangzó szövegek generálásakor, hanem az is, hogy felhívjuk a figyelmet arra, milyen szempontok mentén érdemes tesztelni ezeket a hanggeneráló alkalmazásokat a jövőben, ha természetes hangmintákat szeretnénk kapni.

A konkrét szöveget felolvastató hanganyagok készítéséhez négy alkalmazást teszteltünk: az MS Word felolvasó funkcióját, a ChatGPT hangalapú csevegőfunkcióját, az Elevenlabs kimondottan hanganyaggenerálásra és hangklónozásra optimalizált felületét és a Suno adott szövegre mesterséges zenét előállító alkalmazását. A vizsgált aspektusok a magyar szegmentális és szupraszegmentális jelenségek voltak, beleértve a hezitálási mintázatokat és a beékelten idegen szavak ejtését. Az eredmények alapján az Elevenlabs felület teljesített a legjobban, a ChatGPT és a Suno alkalmazásokban esetenként még akcentusos ejtést tapasztalhatunk (az angol, illetve spanyol nyelv hatását). Egyik alkalmazás sem volt alkalmas az eldöntendő kérdések intonációjának konzisztens reprodukálására, illetve a közbeékelten idegen (angol vagy latin) szavak megfelelő kiejtésére.

Összességében az első vizsgálat eredménye az, hogy az alkalmazások az intonáció és a beékelten idegen szavak kiejtésének kivételével – amely aspektusokat még nem kezelnek jól – változó minőségben teljesítenek, azonban az Elevenlabs a legsikeresebb. Mivel ebben a programban jelenleg nagyon kevés a beépített, ingyenesen elérhető magyar nyelvű hangminta, a későbbiekben választható minták számának emelkedésével vélhetően jobb minőségre számíthatunk majd. Az alkalmazáson belül továbbá lehetőség van hangklónozásra is, így saját hangmintákat is létrehozhatunk. A beékelten idegen szavak és a ritkább szavak megfelelő kiejtése különösen fontos az orvosi szaknyelv esetén, hiszen jellemzően nagyszámú ilyen jellegű terminussal találkozhatunk nem csak szakszövegekben, de egyszerűbb betegdokumentációban is. Ezért az MS Word alkalmazás használata mesterséges hangminták létrehozására szaknyelvoktatásban jelenlegi állapotában nem javasolt.

Második vizsgálatunkban az autonóm tanulásra is alkalmazható chatbotok terén a Mizou, a Gliglish és a ChatGPT alkalmazásokat teszteltük, valamint a Notebook LM podcastgeneráló funkcióját. Ezek az alkalmazások nem általunk bevitt konkrét szövegeket hangosítottak fel, hanem a chatbotok esetén megadott orvos–beteg párbeszéd szituációra reagáltak, míg a Notebook LM egy tananyag alapján készített podcastot, így az eredmény kevésbé volt megjósolható és az output elicíthető. Ezért az első vizsgálatához képest kevesebb jelenséget elemeztünk a kiejtésben (szegmentális és szuprasegmentális jegyek, beszédtempó), azonban ezeken túl tartalmi-nyelvi szempontokat is figyelembe vettünk: a tegezés/magázás problémakörét, a használt releváns szakszavakat (szakzsargon vagy betegbarát megfogalmazások), illetve az orvos részéről a kapcsolattartó fatikus elemek jelenlétét. A kiejtés kapcsán elmondható, hogy a chatbotokban hiányosságként merült fel az eldöntendő kérdések kérdő intonációjának reprodukálása, a Notebook LM pedig még számos kiejtésbeli hibát tartalmazott a magyar podcast hanganyagában. Az orvos–beteg párbeszéd szimulálásakor a botok közül csak a ChatGPT magázta a beteget, a szakszavak tekintetében viszont releváns megoldásokat kaptunk, inkább a laikus regiszterből. A kapcsolattartó fatikus elemek jellemzően minden megnyilatkozás elején szerepeltek, de a ChatGPT esetén semmilyen változatosság nem volt ezek terén tapasztalható.

Összegzésként azt láthattuk, a tesztelt alkalmazások kiejtés tekintetében még nem tökéletesek. A jövő kihívásai között érdemes például az eldöntendő kérdések hanglejtését említeni, mint gyenge pontot. A technológia rohamos fejlődésével azonban várhatóan hamarosan rendelkezésünkre állnak majd a magyar mint idegen nyelv tanítása esetében is olyan alkalmazások, amelyek megbízható oktatói asszisztensként a jelenlegi hiányosságok kiküszöbölésével már az elvárt színvonalon teljesítenek.

Vizsgálatunk korlátai között említhető a pilot jelleg, hiszen relatíve alacsony számú mintán, így kevés adattal végeztük, illetve az a tény, hogy 2025-ben került rá sor. 2026-ban jelentős újítások érhetők el több platform esetén is: az Elevenlabs felületén már jóval több magyar nyelvű hang érhető el, érzelmeket is képesek megjeleníteni a mintákban, ha a promptoló megfelelő annotációkkal látja el a felolvasatni kívánt szöveget, illetve dialógusok is létrehozhatók. A Notebook LM magyar podcastot és videót is képes generálni már, szinte tökéletes (de még mindig nem hibátlan) magyar kiejtéssel. Szakmai tartalom szempontjából azonban megbízhatóbb, ha hivatalos orvosi chatbotokat használunk, mint például a jelenleg már nem ingyenes Magyar Orvosi ChatGPT Pro (Pukoli, 2024) változatát.

Hivatkozások

- Baditzné Pálvölgyi, K. (2020): Patrones de titubeo en el habla espontánea de estudiantes de ELE húngaros: la influencia de la inmersión en el país meta. *Revista Horizontes de Lingüística Aplicada*. 19/1. 145–158. <https://doi.org/10.26512/rhla.v19i1.26945>

- Baditzné Pálvölgyi, K. (2024): Tananyagfejlesztés kezdő spanyol orvosi szaknyelv órákon a mesterséges intelligencia segítségével. *Porta Lingua* 2024/2. 135–142. <https://doi.org/10.48040/PL.2024.2.11>
- Barrington, S. – Cooper, E.A. – Farid, H. (2025): People are poorly equipped to detect AI-powered voice clones. *Scientific Reports*. 15. 11004 (2025). <https://doi.org/10.1038/s41598-025-94170-3>
- Chan, C. – Kuang, J. (2025): Toward Objective and Interpretable Prosody Evaluation in Text-to-Speech: A Linguistically Motivated Approach. arXiv preprint arXiv:2511.02104. <https://doi.org/10.48550/arXiv.2511.02104>
- Chiu, J. – Castro, B. – Ballard, I. et al. (2025): Exploration of the Role of ChatGPT in Teaching Communication Skills for Medical Students: A Pilot Study. *Medical Science Educator*. 35. 1871–1882. <https://doi.org/10.1007/s40670-025-02394-9>
- Goldman-Eisler, F. (1968): *Psycholinguistics: Experiments in Spontaneous Speech*. London: Academic Press.
- Gósy, M. (1988): Tempóészlelés és beszédmegértés. *Műhelymunkák V.* 87–122.
- Gósy, M. (2004): *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.
- Harris, P. T. S. (2024): *Faculty Perspectives Toward Artificial Intelligence in Higher Education*. A Research Paper Submitted to the School of Computing Faculty Middle Georgia State University. Macon, Georgia.
- Holec, H. (1981): *Autonomy and foreign language learning*. Oxford: Pergamon.
- Horváth, L. (2023): Mesterséges Intelligencia alkalmazási lehetőségei a tanítási-tanulási folyamatban. *Plenáris előadás. A tanulás jövője – Együtt a fejlődés útján konferencia*. 2023. december. 6.
- Kassai, I. (1993): Gyorsult-e a magyar beszéd tempója az elmúlt 100–120 évben? In: Gósy, M. – Siptár, P. (szerk.) *Beszéd kutatás '93. Tanulmányok az elméleti és az alkalmazott fonetika köréből*. 62–69. Budapest: MTA Nyelvtudományi Intézet
- Kohnke, L. – Moorhouse, B. L. – Zou, D. (2023): ChatGPT for Language Teaching and Learning. *RELC Journal*. 54/2. 537–550. <https://doi.org/10.1177/00336882231162868>
- Kuna, Á. (2024): *Az orvos–beteg kommunikáció nyelvészeti elemzése. Elmélet, gyakorlat, módszer*. ELTE Eötvös Kiadó: Budapest. <https://doi.org/10.21862/BeszKutAlk2024/Kuna/6944>
- Li, Y. A., Han, C. – Jiang, X. – Mesgarani, N. (2023): Phoneme-level bert for enhanced prosody of text-to-speech with grapheme predictions. In: *ICASSP 2023–2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1–5. IEEE. <https://doi.org/10.1109/ICASSP49357.2023.10097074>
- Little, D. (1991): *Learner autonomy: Definitions, issues and problems*. Dublin: Authentik.
- Orrego Luzoro, M. – Morán Córdova, N. (2014): *Ortopedia y traumatología básica*. Universidad de los Andes: Santiago de Chile.
- Rajki, Z., T. – Nagy, J., – Dringó-Horváth, I. (2024): A mesterséges intelligencia a felsőoktatásban: – hallgatói hozzáférés, attitűd és felhasználási gyakorlat. *Iskolakultúra*. 34/7. 3–22. <https://doi.org/10.14232/iskult.2024.7.3>
- Varga, L. (1994): A hanglejtés. In: Kiefer, F. (ed.). *Strukturális Magyar Nyelvtan*, 2, Fonológia. Budapest: Akadémiai Kiadó. 468–549. (2. kiadás: 2001.). <https://doi.org/10.1556/aling.49.2002.3-4.4>
- Wang, S. – Székely, É. (2024): Evaluating text-to-speech synthesis from a large discrete token-based speech language model. In: *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*. 6464–6474. <https://doi.org/10.63317/22fg5iyuwxfe>

Internetes hivatkozások

- Amithasagaran, A. – Dakshit, S. – Suryadevara, B. – Stockton, L. (2025): *CLiVR: Conversational learning system in virtual reality with AI-powered patients* (arXiv:2510.19031). Online elérhető: <https://arxiv.org/abs/2510.19031>. <https://doi.org/10.48550/arXiv.2510.19031>
- Freeman, J. (2024): *Provide or punish? Students' views on generative AI in higher education*. Higher Education Policy Institute. Online elérhető: <https://www.hepi.ac.uk/wp-content/uploads/2024/01/HEPI-Policy-Note-51.pdf>
- Karageorgakis, T. (2023): Best 100 AI apps for Teachers and the classroom. *Educraft.tech*. Online elérhető: <https://educraft.tech/best-100-ai-apps-for-teachers>
- NielsenIQ. (2024): *NIQ research uncovers hidden consumer attitudes toward AI-generated ads: Negative brand halo and perception effects* [Press release]. Business Wire. Online elérhető: <https://www.businesswire.com/news/home/20241212354811/en/NIQ-Research-Uncovers-Hidden-Consumer-Attitudes-Toward-AI-Generated-Ads>
- Pukoli, D. (2024): Mesterséges intelligencia az orvostudományban: a Magyar Orvosi ChatGPT működése. *MOK. Orvosok Lapja*. 2024. 11. 11. Online elérhető: <https://mok.hu/orvosoklapja/aktualis/mesterseges-intelligencia-az-orvostudomanyban-a-magyar-orvosi-chatgpt-mukodesee>

- Robert, J. (2024): *The future of AI in higher education* (2024 EDUCAUSE AI Landscape Study). EDUCAUSE. Online elérhető: <https://www.educause.edu/ecar/research-publications/2024/2024-educause-ai-landscape-study/the-future-of-ai-in-higher-education>
- Schroeder, R. (2024): How Will AI Disrupt Higher Education in 2024? *Inside Higher Education*. Online elérhető: <https://www.insidehighered.com/opinion/blogs/online-trending-now/2024/01/05/how-will-ai-disrupt-higher-education-2024>
- Stannard, R. (2024): 5 AI Tools to Study Languages At Home- Multiple Languages. *Teacher Training Videos-YouTube Channel*. 2024. 05. 20. Online elérhető: <https://www.youtube.com/watch?v=C6A9quEJETk>

Köszönetnyilvánítás

Köszönettel tartozom az ELTE Digitális Oktatásfejlesztési Kompetencia Központ 2022-1.1.1-KK-2022-00003: „Mesterséges intelligencia az oktatásban labor (MIOL)” pályázatának a kutatás támogatásáért.

An investigation of AI-generated audio materials in the teaching of Hungarian medical language

This research was conducted with the aim of providing a concise overview for instructors of Hungarian as a foreign language regarding AI-based tools—some of which are freely accessible—that can efficiently support and facilitate language teachers’ work through the creation of audio materials. We primarily examine the currently available options from the perspective of medical language instruction; however, the findings are also relevant for professionals teaching Hungarian as a foreign language at any level and in any context. The study investigated how AI performs in the following areas in the context of Hungarian medical terminology: audio materials created by instructors from specialized texts, as well as chatbots capable of autonomously generating audio content or even exchanging voice messages, optimized specifically for medical contexts and autonomous learning. Based on the results, voice generation—including samples created through voice cloning—remains an area in need of development. Applications are currently less capable of producing natural speech intonation. Overall, when examining segmental and suprasegmental features, ElevenLabs stood out among similar applications. Chatbots intended for autonomous language learning can also only be used cautiously at present; there are still shortcomings in pronunciation, while the vocabulary tends to be patient-centered and in a lay register.

Keywords: Hungarian medical terminology, LSP teaching, synthetic speech, autonomous learning, chatbots