

Az izolált mandarin kínai lexikai tónusok akusztikai elemzése kínaiul tanuló magyar anyanyelvűek ejtésében

Juhász Kornélia

*Nyelvtudományi Kutatóközpont
ELTE Eötvös Loránd Tudományegyetem
MTA-HUN-REN NYTK Neurofonetikai Kutatócsoport*

Abstract

The experiment aims to provide an acoustic phonetic investigation into how Hungarian learners of Mandarin produce isolated Chinese lexical tones. In Mandarin Chinese (MC), four lexical tones are contrasted: high level Tone 1 (T1), rising Tone 2 (T2), low falling-rising Tone 3 (T3), and falling Tone 4 (T4). These four tones do not exclusively differ by their F0 curve, but their duration also serves as an acoustic cue for differentiation. The primary focus of the study is to acoustically compare L2 learners and MC natives' production by two acoustic characteristics: the duration, as well as the shape of the f0 curves. The speech of two L2 learner groups (beginners, advanced learners) was compared to a native MC control group (8 speakers per group, all women). Speakers were asked to read CV-structured meaningful words (*ma* syllables), characterized by the four lexical tones. The analysis included the comparison of the duration of the syllable, as well as contrasting the individual tonal realizations' f0 contours among the three speaker groups. The results show that both L2 learner groups produced Mandarin lexical tones with the same temporal characteristics as native speakers, both for absolute and relative durations. Concerning the shape and register of the f0 curves, both L2 learner groups produced the isolated tonal patterns similarly to native MC speakers, i.e., the production of the four lexical tones did not pose problems for Hungarian learners. Some minor differences were observed primarily in the case of beginners, whose production differed more from the native patterns compared to advanced learners: the concavity and the f0 range of the T2, T3, and T4 were distinct compared to the MC native realizations. The significance of the study is that, to the author's knowledge, it is the first analysis that provides statistically validated results on the acoustic comparison of isolated lexical tone production in the speech of Hungarian learners of MC.

1. Bevezetés

A jelen kísérlet célja az izolált ejtésű mandarin kínai lexikai tónusok produkciójának átfogó akusztikai fonetikai vizsgálata kínaiul tanuló magyar anyanyelvű

Email address: juhasz.kornelia@nytud.elte.hu (Juhász Kornélia)

beszélők ejtésében. A tanulmányban a négy kínai lexikai tónust két akusztikai szempont szerint, vagyis a dallamívük menete, valamint a megvalósulásuk időtartama szerint hasonlítom össze két, eltérő mértékű célnyelvi tapasztalattal rendelkező kínaiul tanuló csoport ejtésében. A tanulmány jelentősége az, hogy a szerző tudomása szerint először nyújt statisztikai módszerekkel megerősített eredményeket a kínaiul tanuló magyar anyanyelvűek produkciójáról olyan tekintetben, hogy az elemzés mind a négy lexikai tónus izolált ejtésére is kiterjed. A következőkben bemutatom, hogy hogyan valósul meg a beszéddallam kihasználása a szóban forgó, a nyelvtanulók anyanyelveként (L1) megjelenő magyar, illetve az elsajátítandó célnyelvként (L2) megjelenő mandarin kínai esetében.

Az atonális, avagy monoton (tehát lexikai tónusokkal nem rendelkező) nyelvek esetében, mint amilyen a magyar is a beszéd hangmagasságának változása elsősorban mondat- vagy közlésszintű egységek jelentését változtatja meg (Markó, 2017). Tonális (politon) nyelvek esetében (mint amilyen a mandarin kínai) azonban a hangmagasságváltozás elsősorban lexikai szinten határozza meg a jelentést: a lexikai tónus a szó(tag) argumentuma. Ebből következően ugyanahhoz a szó(tag)hoz eltérő hangmagasságváltozás-mintázatok (dallamkontúrok) társulhatnak, amelyek jelentésmegkülönböztető szereppel bírnak (Chao, 1948/1963). Habár a kínai nyelv tonális, ez nem jelenti azt, hogy a kínai anyanyelvűek beszédében a mondatok/közlések szintjén ne jelenne meg az intonáció: a lexikai tónusok akusztikai megvalósulását az intonáció formálja és határozza meg (Shen, 1989). Akár atonális, akár tonális nyelvről van szó, a beszédképzésben mind az intonáció, mind a lexikai tónus létrehozása ugyanazzal a fiziológiai jelenséggel, azaz a hangszalagműködés változásával áll összefüggésben. A hangszalagok nyitódásából és záródásából előálló kváziperiodikus rezgést zöngének nevezünk. A zöngé egyik alapvető jellemzője az alaphfrekvenciája (f_0), amely az egy időegység alatt lezajlott periódusok számát jelenti. Az intonáció, illetve a tónusok produkciójához szükséges változás az alaphfrekvencia változásából fakad, ugyanis a beszédhang észlelt hangmagassága az alaphfrekvenciával függ össze úgy, hogy a magasabb alaphfrekvencia magasabb hangmagasság érzetét kelti ('t Hart et al., 1990; Gósy, 2004). Abból fakadóan, hogy a beszédhangmagasság és az f_0 -értékek

közötti összefüggés nem lineáris, hanem inkább logaritmikusan tekinthető, a beszélők közötti eltérések normalizálása végett az f_0 -értékeket félhangokká szokás konvertálni (vö. Nolan, 2003, illetve e jelenség részletes bemutatását lásd: Juhász, 2023). A tanulmányban az alapfrekvencia-értékeket – a más szerzőktől hivatkozott ábrákat kivéve – minden esetben félhangokként jelenítem meg. A félhangok használatával kapcsolatban érdemes azt is megemlíteni, hogy a félhang-távolságokat lehetőség van zenei hangközökben kifejezni és számszerűsíteni, megkönnyítve a lexikai tónusok f_0 -terjedelmének bemutatását (vö. pl. Bolla, 1995).

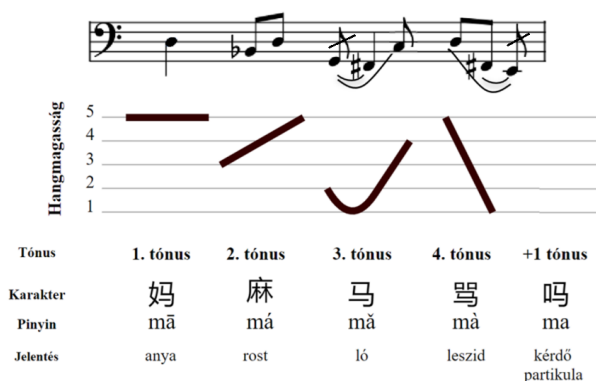
A mandarin kínai nyelv, más szóval a sztenderd kínai nyelvváltozat a hét nagy kínai dialektuscsoport közül az észak-mandarinon, elsődlegesen a pekingi változaton alapszik. A mandarin kínaiiban 4(+1) lexikai tónust tudunk elkülöníteni: a magas szinttartó 1. tónust (T1), az emelkedő 2. tónust (T2), az ereszkedő-emelkedő 3. tónust (T3) és az ereszkedő 4. tónust (T4). Ezen tónusok dallammeneteit az 1. ábra mutatja be. Az 1. ábrán a négy lexikai tónus mellett megjelenő +1 tónust neutrális tónusnak szokás megnevezni, és a neutrális tónusú morfémákat tónustalan szótagnak szokás tekinteni, hiszen e morfémák nem rendelkeznek önálló dallammenettel, a megvalósulásuk a tónuskörnyezet és az intonáció függvényében változik (Shen, 1989; Juhász, 2024). A jelen tanulmány kizárólag a négy teljes értékű lexikai tónusra fókuszál, és a neutrális tónusok ejtésével – rendkívül diverz megvalósulásukból fakadóan – nem foglalkozik. Az 1. ábrán látható kottáról az olvasható le, hogy egy férfi beszélő ejtésében a szinttartó T1 relatíve magas hangmagasság-tartományban, stabilan kitartva (D hangként) valósul meg, Ugyanezt a hangmagasságot (a D hangot) jeleníti meg az emelkedő T2 záró fázisához tartozó hangjegy, valamint az ereszkedő T4 tónus ereszkedő fázisa. Míg az emelkedő T2 esetében egy négy félhangos (nagy tercnyi) emelkedést figyelhetünk meg a kottaképben ($Bé-D$), addig az ereszkedő T4 esetében egy 10 félhangnyi (kis szeptim) ereszkedés figyelhető meg ($D-E$). Az ereszkedő-emelkedő T3 esetében pedig egy kis szekundnyi (1 félhangnyi, $G-Fisz$) lelépést egy bővített kvartnyi (hat félhangnyi, $Fisz-C$) fellépés követ. Az ereszkedő-emelkedő T3 ejtése esetében mindenképpen érdemes megemlíteni azt

is, hogy a pekingi dialektusban a T3 alacsony f_0 -tartományban megvalósuló ejtése irreguláris zöngéképzéssel, pontosabban recsegő/rekedtes fonációval (creaky voice) társul. Ez azt jelenti, hogy a T3 esetében az alapfrekvencia lelassul és a hangszalagok feszítettsége is csökken (Markó, 2013). E képzési tulajdonság alapvetően korlátozza e lexikai tónus akusztikai vizsgálatát abból fakadóan, hogy a legtöbb akusztikai elemzésben is használt szoftver (például a jelen tanulmányban is használt Praat (Boersma & Weenink, 2022) az irreguláris zöngé frekvenciatartományában már nem mér megbízhatóan, illetve egyáltalán nem mér f_0 -értékeket (Dorreen, 2017; Dallaston & Docherty, 2019).

Visszatérve a lexikai tónusok akusztikai tulajdonságaira és megjelenítésére, ezen kottaképekkel bizonyos szempontból ellentmond az öt fokozatú hangmagasságskála és a lexikai tónusok ebben a paradigmában való bemutatása. Az öt fokozatú hangmagasságskálán a szinttartó T1 tónust 5–5-ként szokás kódolni, az emelkedő T2 tónust 3–5-ként, az ereszkedő-emelkedő T3 tónust 2–1–4-ként és az ereszkedő T4 tónust 5–1-ként (Chao, 1948/1963). Fontos kiemelni, hogy ezen sematizált öt fokozatú hangmagasságskála szolgál a kínai lexikai tónusok elsjátításakor az elsődleges referenciának. Habár a skálán bemutatott mintázatok jól bemutatják a hangmagasság-változás irányát és relatív pozícióját, azonban például a minimális f_0 -érték szempontjából elfedik az ereszkedő-emelkedő T3 és az ereszkedő T4 tónus közötti kottaképen bemutatott eltérést. Ha az 1. ábrán látottakkal összevetjük a tónuskontúrok tényleges megvalósulását, vagyis az f_0 -kontúrjaikat (2. ábra), akkor még több inkonzisztenciát fedezhetünk fel az egyes megjelenítések között, elsősorban a maximális f_0 -értékek tekintetében. Az 1. ábra kottája és a sematikus hangmagasságskála egyöntetűen egyezőként mutatja be a T1, T2, illetve a T4 maximális hangmagasságát. Ehhez képest az f_0 -kontúrok maximuma ennél sokkal árnyaltabb: e három felsorolt lexikai tónus f_0 -kontúrbeli realizációja közül a T4 rendelkezik a legmagasabb, a T1 pedig a legalacsonyabb f_0 -maximummal, míg a T2 maximuma éppen e két tónus között helyezkedik el.

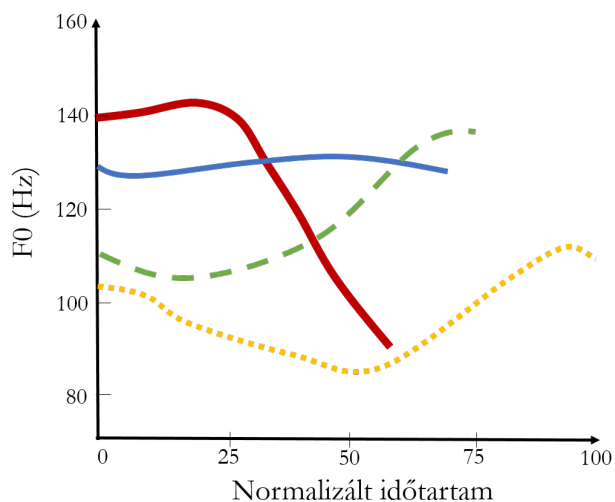
A lexikai tónusok esetében a dallamkontúr mellett egy másik fontos akusztikai tulajdonságukat is meg kell említeni: a lexikai tónusok időtartamának

relevanciáját. Zhang és munkatársai (2022) eredményei alapján a kínai suttogott beszédben az időtartam a tónus másodlagos felismerési kulcsaként szolgál, és a kísérletben a kínai anyanyelvű beszélők hajlamosak voltak kihangsúlyozni a lexikai tónusok közötti időtartambeli különbségeket suttogott beszédben azért, hogy a suttogásban neutralizálódó, a dallam által kifejezett kontrasztot fenntartsák (Zhang et al., 2022). Az időtartambeli különbségek a 2. ábrán is jól megfigyelhetők: az f₀-kontúrok időtartama a leghosszabb tónushoz, a T3-hoz van normalizálva. Ehhez képest a T2 időtartamában csak $\frac{3}{4}$ olyan hosszú, mint a T3, hasonlóan a T1 megvalósulásához (amely $\sim 70\%$ -a a T3 időtartamának). A négy tónus közül az ereszkedő T4 a legrövidebb, amely a T3-hoz képest majdnem fele olyan hosszú ($\sim 55\%$ -a a T3-nak).



1. ábra. A tónusok dallamának kottaképe férfiak ejtésében (felső sor) és sematikus dallamkontúrjai (középső sor) (Chao, 1948/1963: 85 nyomán), valamint az adott tónussal megvalósuló kínai szótagok (kínai karakterekkel, illetve a hangjelölő pinyin átírással megjelenítve, valamint a szavak jelentése) (Juhász, 2024 nyomán, javítva).

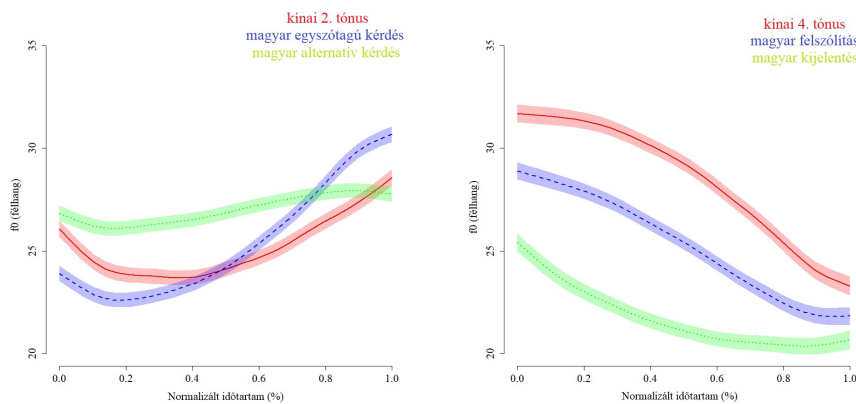
Az idegennyelv-elsajátítás (L2/cél nyelv-elsajátítás), vagyis a jelen esetben a kínai lexikai tónusok ejtése és elsajátítása több szempontból is problémát okozhat az atonális magyar anyanyelvűeknek. Ha a dallamkontúrok jelentésközvetítő funkcióját vesszük középpontba, akkor az atonális magyar nyelvben is természetesen megfigyelhetők – a kínai tónusokhoz hasonlóan – monoszillabikus egységeken megvalósuló dallamok. Egy korábbi kísérletben e dallam-realizációk



2. ábra. A kínai lexikai tónusok f₀-görbéje a T3 időtartamához normalizálva, ahol T1 = vékony folytonos kék vonal, T2 = szaggatott zöld vonal, T3 = pontozott sárga vonal, T4 = vastagított folytonos piros vonal (Xu, 1997: 67 nyomán).

összehasonlítása végett az emelkedő és ereszkedő kínai lexikai tónusokat egy szótagú magyar intonációs dallamokkal is összevettem. E kísérlet alapját az a kínai tanítási gyakorlatban gyakran megjelenő módszertan adta, miszerint a magyar anyanyelvű kínai tanárok az emelkedő T2 tónust a magyar egy szótagú kérdés dallamával, míg az ereszkedő T4 tónus a megszólítás/felszólítás dallamával feltételezik egyezőnek, és a tónuselsajátítást is ilyen módon instruálják. A kínai tónusok és a magyar intonációs kontúrok megfeleltetését a következő akusztikai összevetésben vizsgáltam: az emelkedő T2 tónust a magyar egy szótagú eldöntendő kérdő dallammal (pl. *Én?*), valamint az alternatív kérdés első szótagjának emelkedő dallamával vettem össze (lásd a kiemelt szót a következő megnyilatkozásban: *Én vagy ő?*). Az ereszkedő T4 tónust a magyar egy szótagú vokatív megszólítás és a kijelentés ereszkedésével vettem össze. A kísérlet eredményei rámutattak arra, hogy e két kínai tónus f₀-kontúrjának lefutása minden páros összehasonlításban eltér a velük összevetett magyar dallamok megvalósulásától. A kínai emelkedő tónus a magyar egy szótagú kérdő, illetve az alternatív

dallamhoz képest homorúbban valósult meg, és az egy szótagú kérdésnél kisebb, míg az alternatív kérdés emelkedésénél nagyobb f_0 -terjedelem jellemezte (3. ábra bal). Az ereszkedő T4 tónus mind a magyar vokatívusz, mind a kijelentésnél domborúbb dallamívvel rendelkezett (3. ábra jobb) (Juhász, 2023). A kínai oktatásban a kínai szinttartó T1 tónus és az ereszkedő-emelkedő T3 tónus esetében nem szokás a magyar intonációval, mondattípusokkal kapcsolatos párhuzamot vonni, valószínűsíthetően azért, mert nincs olyan magyar dallamkontúr, amely egyértelműen és kizárólagosan példaként szolgálhatna az említett kínai tónusok esetében.



3. ábra. A kínai T2 és a magyar egy szótagú kérdő dallam, valamint az alternatív kérdő dallam emelkedő dallamenete (balra), valamint a kínai T4 és a magyar egy szótagú vokatívusz és kijelentés f_0 -kontúrjai (jobbra) (Juhász, 2023: 35, 38).

A fentebb bemutatott eredmények az idegennyelv-elsajátítás középponti problémáját alapozzák meg: amennyiben egy új tanulási szituációval kell megküzdenünk, hajlamosak vagyunk a már elsajátított és ismert megoldási mintázatokat alkalmazni a problémamegoldás során. Az idegennyelv-elsajátításban ezt a problémamegoldási stratégiát transzfernek nevezzük (Odlin, 1989; Major, 2001). A transzfer-hipotézis alapján a jelen esetben az feltételezhető, hogy az idegen nyelvi kínai lexikai tónusok elsajátításakor a nyelvtanuló a már korábban elsajátított, többnyire anyanyelvi dallamkontúrait alkalmazza a produkciójában. Azt fontos kiemelni, hogy a nyelvek közötti transzfer nem korlátozódik az anya-

nyelv célnyelvre gyakorolt hatására, hiszen a nyelvtanuló bármelyik korábban tanult idegen nyelvei is formálhatják a produkciót (Westergaard et al., 2017). Azonban a kísérletben vizsgált nyelvtanulók legdominánsabb nyelve kétségkívül a magyar, vagyis az anyanyelv, így az L1-es mintázatok tekinthetők a transzferjelenség elsődleges forrásának, (főleg abban az esetben, ha ezt a tanítási környezet is facilitálja). Így tehát az eredmények esetében potenciálisan megjelenhet az anyanyelv hatása a lexikai tónusok megvalósulásában. Fontos megemlíteni ugyanakkor, hogy az anyanyelvi hatás közvetlen vizsgálatát kizárólag úgy lehetne megerősíteni, ha a kínaiul tanuló magyar anyanyelvűek lexikaitónus-produkcióját a hozzájuk leghasonlóbb magyar monoszillabikus intonációs dallamok megvalósulásával is összehasonlítjuk, azonban erre a jelen tanulmány nem terjed ki.

A szakirodalmi forrásokban az idegennyelv-elsajátítás folyamatát, valamint az akcentusjelenségeket kiváltó változókat mind a percepció, mind a produkció aspektusából is szokás tekinteni. Az L2-elsajátítási modellek többsége a percepció elsődlegességét feltételezi a produkcióhoz képest, hiszen az emberi észlelés révén jutunk folyamatos visszajelzéshez az ejtett beszédhangok/dallamok megvalósulásáról, és ezáltal tudjuk formálni a produkciónkat (Kuhl, 1980). Ugyanakkor például Flege és Bohn (2021) Revised speech learning model-je a percepció és a produkció szimultán fejlődését feltételezi az idegennyelv-elsajátításban (az L2-elsajátítási modellek részletes bemutatását lásd: Juhász, 2024). A tanulmány szempontjából a kínai lexikai tónusok kínai natív mintától eltérő, akcentusos ejtése mind a percepció, mind a produkció hatásából is bekövetkezhet (vö. Klein, 1986). A percepció akcentus alapja az, hogy az emberi észlelésben az L1 rendszere egyfajta fonológiai szűrőként működik, és ezáltal, ha az anyanyelvi és célnyelvi mintázatok (legyenek ezek dallamkontúrok vagy éppen beszédhangok) közötti fonetikai eltérés nem jelentős, akkor az anyanyelv szempontjából nem kontrasztív akusztikai különbségek kiszűrődhetnek a percepcióból (Flege, 1995; Flege & Bohn, 2021). Például ha a fentebb bemutatott T2-es tónust és a magyar egy szótagú kérdés dallamát tekintjük, Mennen az idegen nyelvi intonáció elsajátítására fókuszáló modellje (2014) alapján azt állíthatjuk, hogy ezek

nem csak elemezhető az autoszegmentális fonológia paradigmáján belül, hanem ugyanazon mögöttes alacsony-magas (LH) fonológiai célokkal írhatók le, amelyek emelkedő dallammenetként realizálódnak a produkcióban. Azonban ezen emelkedő dallammenetek, mint láttuk, az emelkedő dallammenet ellenére fonetikailag számos eltérő akusztikai jeggyel rendelkezhetnek (pl. eltérő f₀-tartomány, maximális f₀, minimális f₀, a görbe meredeksége, domborúsága stb.). Az anyanyelvi perceptuális szűrő fonológiai természetéből következően funkcionális elvek szerint működik, vagyis ha az anyanyelvben például az f₀-kontúr domborúsága nem rendelkezik jelentésmegkülönböztető szereppel, akkor feltételezhetően ezen akusztikai jegyben bekövetkező eltérés észlelése – a nyelvelsajátítás legelejét tekintve legalábbis – nem jelenik meg a percepcióban. E folyamat eredményeképpen a célnyelvi mintázatot a nyelvtanulók könnyen az anyanyelviként egyezőnek észlelhetik, és transzferálhatják produkciójukba is. Ha a nyelvtanuló az L1 és L2 között fennálló akusztikai különbségek dacára a célnyelvi mintázatot az anyanyelvi megfelelőjével helyettesíti, azt Flege (1995), valamint Flege és Bohn (2021) munkája alapján ekvivalensként való osztályozásnak nevezzük. Ugyanakkor az akcentusos ejtés problematikája produkciós nehézségekből is eredhet: ebben az esetben a nyelvtanuló észleli az L1 és L2 mintázata közötti különbséget, pusztán a dallamkontúrok akusztikai megvalósítása nem közelíti meg a natív mintázatot. Ebben az esetben például az anyanyelvi artikulációs-motorikus rutinok célnyelvi mintázattal való felülírása is megnehezítheti az akcentusmentes ejtés elsajátítását (Leather & James, 1991; Mennen, 2014). Ezen felül a dallammenetek elsajátítására még hatással lehet az adott dallamok megjelenési gyakorisága is (Mennen, 2014): abból fakadóan, hogy a jelen tanulmány izolált lexikai tónusokat vizsgál, amelyeket minden kínai nyelvtanuló már az első kínai nyelvórán, megközelítőleg azonos mennyiségben észlel és produkál, ebben a tekintetben a dallamok megjelenési gyakorisága feltételezhetően nincs közvetlen hatással az ejtésre (szemben például bizonyos tónuskapcsolatokkal, ahol a különböző tónusszekvenciák megjelenési gyakorisága szóródhat). Ehhez kapcsolódóan még azt is meg kell említeni, hogy a tónuskapcsolatok esetében koartikulációs hatásokkal is számolnunk kell, vagyis a tónus-környezet hatására a lexikai tó-

nus akusztikai megvalósulása minimális fonetikai változásokon mehet keresztül (vö. Xu, 1997), amely variabilitás a nyelvtanulók számára nehezítheti a lexikai tónusok ejtését. A szerző saját, mind tanulói, mind oktatói tapasztalatai alapján az izolált ejtésű tónusok elsajátítása (és tanítása) – a tónusszekvenciákhoz képest – jelentős figyelmet kap a kínai oktatásban, amely módszertannak része egyrészt a lexikai tónusok izolált ejtésű és hiperartikulált bemutatása és gyakorlása, hangsúlyozva a különböző tónusokra jellemző és egymástól eltérő akusztikai tulajdonságokat, ami elősegítheti a kínai tónusok közötti kontrasztok kialakítását a nyelvtanuló elméjében. Másrészt a tónuskontúrok akusztikai jegyeinek verbalizálása mellett az f_0 -görbék vizuális megjelenítése (l. pl. 1. ábra) is rendelkezésre áll az akusztikai különbségek nyilvánvalóvá tételére, amely technikák mind elősegíthetik (vagy éppen visszavethetik) a kínai lexikai tónusok natív ejtéshez hasonló produkciójának elsajátítását.

Összegezve, az idegen nyelvi akcentus mind a percepció, mind a produkció szintjén megjelenhet, és e jelenség potenciális felszínre kerülése alapvetően az anyanyelvi hatásból következik, aminek eredményeképpen a kínaiul tanuló magyar anyanyelvűek tónusprodukciójában akár a natív ejtéstől eltérő mintázatokra számíthatunk. Ugyanakkor fontos megjegyezni azt is, hogy – szintén a Flege & Bohn-féle (2021) Revised speech learning model-ből kiindulva – az idegennyelv-elsajátítás egy dinamikus folyamat, amely a nyelvi tapasztalat hatásával együtt és fokozatosan fejlődik. Tehát az anyanyelv mellett a célnyelvi tapasztalat mértéke is egy fontos változó az idegennyelv-elsajátítás szempontjából (Flege & Bohn, 2021), ezért a kísérletben két, eltérő mennyiségű kínai célnyelvi tapasztalattal rendelkező magyar anyanyelvű beszélő ejtését vizsgálom azon predikció tesztelése céljából, hogy vajon a több L2-tapasztalattal rendelkező nyelvtanulók jobban megközelítik-e a kínai anyanyelvi ejtést, mint a kevesebb L2-tapasztalattal rendelkező nyelvtanulók. E tanulmány kizárólag a kínai tónusok ejtésének dinamikus jellemzőit, és a produkció nehézségeit vizsgálja, vagyis az akcentusos ejtés okait nem kívánja empirikus úton feltárni, illetve az anyanyelvi hatás közvetlen vizsgálatára sem tér ki.

Az idegennyelv-elsajátítás alapjairól a kísérlet empirikus előzményeire térve kérdésként merülhet fel az atonális anyanyelvű beszélők hatékonysága a lexikai tónusok tekintetében akár más, tonális anyanyelvű személyekkel szemben. Például Lee, Vakoch és Wurm (1996) eredményei arra mutattak rá, hogy a tonális anyanyelv kizárólag akkor járulhat hozzá egy másik (L2) tonális nyelv lexikai tónusainak elsajátításához (ebben az esetben csak a percepció diszkrimináció tekintetében), amennyiben az anyanyelvi tónusrendszer komplexebb az elsajátítandó nyelvváltozaténál. E tekintetben tehát – a Flege- és Bohn-féle (2021) Revised speech learning modelben megfogalmazottakkal párhuzamosan – az anyanyelvi mintázatok alapján prediktálható az, hogy a nyelvtanulók milyen problémákba ütköznek a lexikai tónusok elsajátításában. Hao (2012) eredményei – bizonyos szempontból ellentmondva az előbbieknél – arra utalnak, hogy habár a kantoniban 6 különböző lexikai tónus jelenik meg, ennek ellenére azonban a kantoni anyanyelvű beszélők mégsem teljesítenek jobban a mandarin tónusok percepciójában és produkciójában az atonális angol anyanyelvű beszélőkhöz képest. Emellett Hao eredményei a mandarin lexikai tónusok elsajátítását is két eltérő nehézségi szintre osztja: mind percepció, mind produkció szempontjából, mind a tonális kantoni, mind az atonális angol beszélők a T1 és T4 lexikai tónusok esetében jobban teljesítettek, mint a T2 és a T3 esetében. A T2 és a T3 esetében tapasztalt percepció és produkció nehézségek ezért feltételezhetően nem elsősorban anyanyelvi hatásnak köszönhetők, sokkal inkább a két lexikai tónus intrinzik akusztikai tulajdonságának, vagyis hasonló lefutásának. Nem beszélve arról az alternációról, hogy T3+T3 szekvencia esetén az első T3 T2-vé változik ($T3+T3 = T2+T3$) (Hao, 2012: 278).

Ha a mandarin lexikai tónusok elsajátítását a kínaiul tanuló magyar anyanyelvűek szempontjából tekintjük, akkor a négy kínai tónus izolált ejtésű megvalósulását korábban Qiuyue Ye (2013) vizsgálta disszertációjában. Az akusztikai elemzésben összesen 12 beszélő vett részt: Ye 10 nyelvtanuló (4 nő és 6 férfi), valamint két natív kínai beszélő (1 férfi és 1 nő) ejtését hasonlította össze kvalitatív minőségben a négy kínai lexikai tónus izolált ejtésű időtartama, illetve f0-kontúrja tekintetében. A nyelvtanulók 9 hónap és 3 év közötti kínai nyelv-

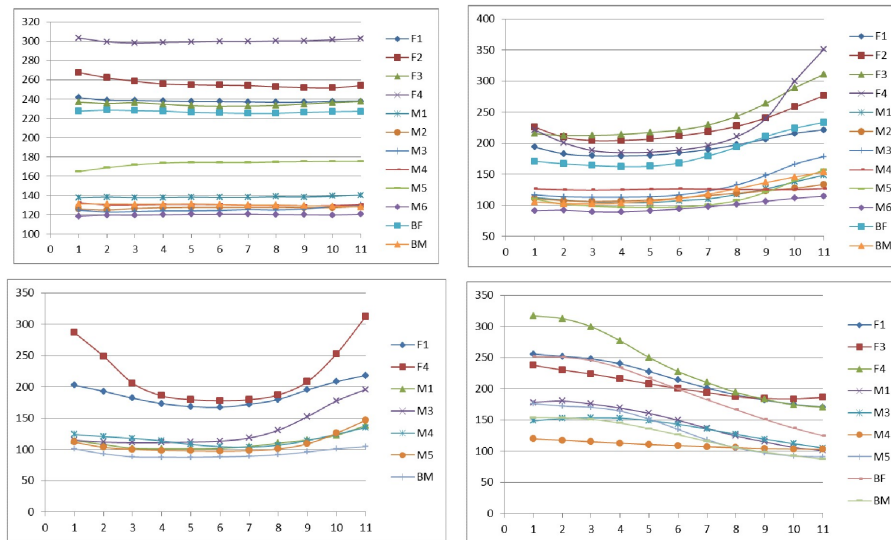
vi tapasztalattal rendelkeztek. Az akusztikai elemzés az egyéni mintázatokra fókuszált, és a nyelvi tapasztalat szóródásából következően csoportszintű produkciót Ye nem közölt, ugyanígy feltételezhetően ugyanebből az okból fakadóan statisztikai elemzést sem mutatott be. Az egyéni mintázatok középpontba helyező eredmények a következőképpen foglalhatók össze, különös figyelmet fordítva a jelen tanulmány szempontjából is fontos akusztikai jellemzőkre, vagyis a lexikai tónusok időtartamára, valamint a f0-kontúrok megvalósulására. Időtartam tekintetében a 10 vizsgált nyelvtanuló jelentős hányada mind a négy tónus esetében a natív megvalósulásoknál rövidebbet produkált (1. táblázat, Ye 2013: 101, 138, 175, 205). Például a nyelvtanulók többsége a T4-et ejtette a többi tónushoz képest a legrövidebb időtartammal, valamint a 10 beszélő közül 8 ejtésében a natív mintának megfelelően a négy tónus közül T3 a leghosszabb. A T1 és T2 esetében a natív mintázat hasonló időtartamokat jelenít meg, azonban a 10 beszélő közül 7 a T2-t rövidebb időtartammal produkálja mint a T1-et (Ye, 2013). Ezen eredmények arra utalnak, hogy a kínaiul tanuló magyar anyanyelvűeknek nem okoz problémát a natív ejtésre jellemző időtartamokkal ejteni a kínai lexikai tónusokat.

1. táblázat. A kínai lexikai tónusok átlagos időtartama (ms) a Ye (2013) disszertációjában vizsgált 12 beszélő ejtésében (Ye, 2013: 101, 138, 175, 205). Az Fx jelölés a magyar női, az Mx jelölés a magyar férfi beszélőket jeleníti meg, míg a BF egy kínai anyanyelvű női, a BM pedig egy kínai anyanyelvű férfi beszélőt jelöl.

	F1	F2	F3	F4	M1	M2	M3	M4	M5	M6	BF	BM
T1	415	492	382	515	283	348	281	352	597	345	421	479
T2	452	462	285	367	313	437	231	330	545	322	451	521
T3	445	574	627	697	426	463	301	481	721	310	641	599
T4	290	312	177	238	207	191	168	267	210	388	288	282

Ye (2013) f0-kontúrokra vonatkozó eredményei beszélők szintjén mutat be átlagértékeket a vokalikus szakaszban 11 ponton mérve és Excelben kirajzolva. Ye (2013) a férfi és női f0-kontúrokat Hertz értékekben mutatja be, és az

ábrákon jól elkülönülnek a férfi és női alaphangtartományok, azonban pusztán az ábrák alapján – a beszélők feltételezhetően eltérő hangszínezetét, illetve a félhangokká konvertálás hiányát is figyelembe véve – meglehetősen nehéz csoportszintű tendenciákra következtetni (4. ábra). Mindenesetre a négy lexikai tónus dallammenetének iránya minden esetben jól megfigyelhető: habár eltérő tartományokban, de a T1 szinttartó f0-mintázata jól kivehető a nyelvtanulók körében, hasonlóan a T2 emelkedésével, de e tónus ejtésében már a görbe meredekségében is látható némi variabilitás a beszélők között, hasonlóan a T3 mély, ereszkedő-emelkedő megvalósulásához. A T4 produkciója minden megjelenített nyelvtanuló esetében ereszkedő mintázatot mutat, azonban kérdésként vetődik fel, hogy a nyelvtanulók a natív ejtésnek megfelelő meredekséggel ejtik-e ezt az ereszkedő dallamot (4. ábra).



4. ábra. Ye (2013) disszertációjában résztvevő beszélők kínai tónuskontúrjai (ahol az f0-értékek a normalizált időtartam 11 pontjában vannak kinyerve és átlagértékeként vannak ábrázolva). A beszélők kódolása: az F betűt tartalmazó kódjelű beszélők nők, az M betűt tartalmazó kódjelű beszélők férfiak és a B-vel kezdődő utolsó két beszélő tekintendő a natív referenciának (Ye, 2013: 126, 163, 194, 224).

1.1. A kísérlet célja és kutatási kérdései

A tanulmányban bemutatott akusztikai elemzés célja az, hogy egy exploratív vizsgálat keretein belül dinamikus, statisztikai módszertannal megerősített eredményeket mutasson be arra vonatkozóan, hogy a kínaiul tanuló magyar anyanyelvűek lexikaitónus-produkciója eltér-e a natív kínai beszélők ejtésétől. Habár a megelőző kutatási eredmények részlegesen lehetővé tehetnék azt, hogy külön-külön hipotéziseket állítsunk fel a négy kínai lexikai tónus ejtésére vonatkozóan, prediktálva az anyanyelvi mintázatok hatására bekövetkező potenciális eltéréseket a kínai anyanyelvű produkcióhoz képest, azonban – a tanulmány exploratív jellegéből fakadóan – pusztán kutatási kérdéseket fogalmazok meg. A hipotézisalkotás elmaradását a következőkkel szeretném indokolni: i.) Amennyiben a kísérlet célja az anyanyelvi hatás vizsgálata, ebben az esetben kizárólag az emelkedő T2-es és az ereszkedő T4-es lexikai tónusok hasonló magyar dallamkontúr-párjairól vannak információink, míg a magas szinttartó T1 és a mély T3 esetében nem áll rendelkezésre statisztikai elemzéssel megerősített mintázat. ii.) Ami a kínaiul tanuló magyar anyanyelvűek kínaitónus-produkcióját illeti, Ye (2013) eredményei kizárólag egyéni mintázatokat mutatnak be, eltérő számú és nemű kínai anyanyelvű és kínaiul tanuló magyar anyanyelvű beszélők ejtésében, mind az f0-kontúrok, mind az időtartam tekintetében, így tehát ezekből az eredményekből csoportszintű ejtésre, és az anyanyelvi mintától való eltérésre nem lehet közvetlenül következtetni. Ezért a jelen tanulmány a következő három kutatási kérdés vizsgálatát tűzi ki célul:

1. A kínaiul tanuló magyar anyanyelvűek a kínai beszélőkkel megegyező időtartammal ejtik a négy mandarin lexikai tónust?
2. A kínai tónusok dallamívét (f0-kontúrját) tekintve a magyar beszélők a kínai natív dallamkontúrokhoz hasonlót produkálnak?
3. További kutatási kérdés az, hogy amennyiben a kínaiul tanuló magyar anyanyelvűek eltérnek a natív kínai mintázatoktól ejtésükben, akkor megerősíthető-e a Flege- és Bohn-féle Revised speech learning model (2021) predikciója,

miszerint a több nyelvi tapasztalattal rendelkező haladó nyelvtanulók jobban megközelítik a natív mintázatokat, mint a kevesebb L2-tapasztalattal rendelkező kezdő nyelvtanulók?

Fontos kérdésként merülhet fel a tónusok rendszerszintű elemzése, vagyis a beszélői csoporton belüli lexikaitónus-kontraszt kérdése, mind az időtartam, mind a dallamkontúrok tekintetében, azonban erre a kérdésre a jelen tanulmány a területi megszorítottságból fakadóan nem terjed ki. A beszélői csoportokon belül a kínai tónusok produkciós elkülönítésére vonatkozó eredményeket egy másik, előkészületben lévő tanulmányban mutatom be (Juhász, előkészületben).

2. Módszertan

A kísérletben két eltérő nyelvi tapasztalattal rendelkező magyar anyanyelvű kínaiul tanuló beszélői csoport ejtését hasonlítottam össze egy kínai anyanyelvű beszélői csoport produkciójával. Mindhárom beszélői csoport 8 kísérleti személyből állt, mind nők. A kezdő kínaiul tanulók kínai alapszakos egyetemisták voltak, akik legalább 1 éve tanulnak mandarinul és nem rendelkeztek Kínában eltöltött ösztöndíjas nyelvi tapasztalattal (átlagéletkoruk $23,2 \pm 3,21$ év volt). A haladó nyelvtanuló csoport tagjai olyan harmadéves kínai szakos hallgatók voltak, akik egy évet Észak- vagy Közép-Kínában töltöttek ösztöndíjjal (4 fő), illetve kínai mesterképzéses hallgatók voltak (4 fő), akik legalább 3 éve tanulnak kínaiul (átlagéletkoruk $24,2 \pm 2,86$ év). A kínai anyanyelvű kontrolleszámítógépes csoport nyolc női beszélője mind északi vagy észak-keleti, mandarin dialektuscsoportokhoz tartozó anyanyelvi beszélő, akik közül négyen Pekingben nőttek fel, négyen pedig Pekingtől maximum 500 km-re (átlagéletkoruk: $23,8 \pm 2,55$ év). A nyolc kínai anyanyelvű beszélő mind egyetemista hallgató, négyük 5 éve, egy beszélő 3 éve, 2 beszélő két éve és egy beszélő 1 éve él Magyarországon és tanul magyar nyelven. Mindannyiuk 15–16 éve tanul angolul, amely nyelvet napi rendszerességgel használnak. A hangfelvételeket 16 bit-en, 44,1 kHz-en digitalizálva rögzítettem egy külső hangkártyával és egy omnidirekcionális kondenzátoros fejmikrofonnal.

A kísérlet anyagát izolált ejtésű CV-szerkezetű, jelentéssel rendelkező, egy szótagú szavak adták. Ezen szótagok egytől egyig a szonoráns *ma* beszédhangokból álltak, ami lehetővé tette a teljes szótagon megvalósuló f0-kontúr vizsgálatát. A *ma* hangsort a négy különböző kínai lexikai tónussal rögzítettem, így a felolvasott szavak a következők voltak: T1: 抹 *mā* (töröl), T2: 麻 *má* (rost), T3: 马 *mǎ* (ló), T4: 骂 *mà* (leszid). A felvételkedéskor mind a kínai karakterek, mind a pinyin transzkripció meg volt jelenítve a kísérleti személyek számára. A kísérleti személyeknek ezen izolált ejtésű szótagokat random sorrendben, 6 ismétléssel kellett felolvasniuk egy képernyőről. Felmerülhet a kérdés, hogy a T1-es 抹 *mā* (töröl) szótag esetében miért ez a karakter került megjelenítésre a sokkal egyszerűbb és gyakrabban használt 妈 *mā* (anya) helyett. Ennek az oka, hogy a jelen vizsgálatban bemutatott izolált tónusprodukción egy olyan akusztikai elemzés referenciájaként fog szolgálni, amelyben a fentebbi négy kínai szótagot egy megnyilatkozás részeként, diszillabikus egységekben is vizsgálom ugyanezen beszélők ejtésében. A 妈 *mā* szótag duplikálásakor azonban a létrejövő 妈妈 *māma* szóban a második tónus neutrális (tónustalan), így nem alkalmas egy T1+T1 konstrukció létrehozására. Megjegyzendő továbbá az, hogy ezen, laboratóriumban rögzített, kísérletvezető jelenlétében elhangzott izolált megnyilatkozások esetében hiperartikulált produkcióra, vagyis az egyes artikulációs mintázatok túléjtésére számíthatunk a spontán beszédben elhangzó megvalósulásokhoz képest (vö. Scarborough & Zellou, 2013), vagyis ebben a szituációban minden körülmény a lexikai tónusok lehetőség szerinti legjobban formált ejtését facilitálja.

A hangfelvételeket a Praat szoftverben (Boersma & Weenink, 2022) címkéztem és elemeztem: minden elemzést a teljes szótagon, vagyis a vokális szakaszon végeztem, és ezt az intervallumot elemeztem a szakasz időtartamként is. Az akusztikai elemzés egy f0-kontúrokat érintő dinamikus és egy tónusidőtartamra fókuszáló statikus vizsgálatot foglal magába.

A statikus időtartamra fókuszáló elemzések esetében vizsgáltam a szótagok abszolút és relatív időtartamát is. Az abszolút időtartam a szótag valós, milliszekundumban kinyert időtartamát jelentette, míg a relatív időtartam esetében

a z -érték mentén beszélőnként normalizáltam az időtartam-értékeket, vagyis vettem az adott időtartamérték és a beszélőre számolt (négy lexikai tónusra számolt) átlag különbségét, majd ezt az értéket elosztottam a beszélőre számolt szórással ($= (\text{érték} - \text{beszélő átlaga}) / \text{beszélő szórása}$). Ez azt jelenti, hogy amennyiben ez a relatív érték nulla körül valósul meg, akkor közel esik a beszélő átlagához, míg például az 1-hez közeli érték arra utal, hogy a beszélő átlagához képest egy szórással hosszabb, míg -1 -es érték esetében egy szórással rövidebb volt az adott megvalósulás. Az abszolút időtartamok helyett a statisztikai elemzésben elsődlegesen a relatív időtartam-értékekre fókuszálok azért, mert ezekben az adatokban az egyes beszélők közötti variabilitás, pontosabban az esetleges beszédtempót érintő eltérések szerepe kevésbé releváns, továbbá ebben az esetben az értékek relativitása a négy tónus egymáshoz való viszonyát is jobban bemutatja (még akkor is, ha a tónus-kontrasztok vizsgálata a jelen tanulmánynak nem célja). A relatív időtartamértékek statisztikai elemzését lineáris kevert modellekkel végeztem, szintén az R programban az lme4 csomag (Bates et al., 2015) segítségével. A p - és F -értékeket Satterthwaite-approximáció segítségével nyertem ki, ami az lmerTest csomagban (Kuznetsova et al., 2017) elérhető. A négy kínai lexikai tónus relatív időtartamát négy különböző lineáris kevert modellel vizsgáltam. Minden tónus esetében a relatív időtartam függő változót vizsgáltam a három beszélői csoport függvényében. Minden tónus esetében egy olyan modellt állítottam fel, amely a csoportok között kizárólag a regressziós egyenes y tengely mentén képzett elmozdulását engedi (a random metszéspont, azaz random intercept, $\text{lmer}(\text{relatív időtartam} \sim \text{beszélői csoport} + (1 | \text{beszélő}))$). Ez összesen négy kevert modellt eredményezett. A függő változó három szintjét (natív/haladó/kezdő) páronként Tukey post hoc tesztekkel vettem össze az emmeans csomag használatával (Lenth, 2020). Az adatokat a ggplot2 csomag segítségével ábrázoltam (Wickham, 2016).

A dinamikus f_0 -kontúrok esetében a 120 Hz felett kinyert f_0 -értékeket minden esetben félhangokká konvertáltam az R programban (R Core Team, 2024) a hqmisc (Quené, 2014) csomag segítségével, minden esetben 50 Hz-es referencia-értékkel. Az f_0 -görbék elemzéséhez az f_0 -értékét a vokalikus szakasz időtartamá-

hoz képest normalizáltan, 1%-onként nyertem ki, automatikusan. Az f0-görbék elemzésére generalizált additív kevert modelleket (GAMM) használtam (Wood, 2017), külön modellel vizsgálva négy lexikai tónus megvalósulását a három vizsgált beszélői csoportban. Ez azt jelenti, hogy a modellekben a vizsgált tónus ejtése a beszélői csoportok között megfigyelhető eltérésekre összpontosított. Az alapmodellben az f0 függő változó alakulását vizsgáltam a normalizált időtartam függvényében, más szóval azt, hogy az f0 értéke az időtartamra simítva hogyan változik a normalizált időtartamon belül. Az f0-görbék girbe-gurbaságából fakadóan az f0-görbék adaptív simítással illesztettem (bs = "ad"), ami lehetővé teszi a görbe hullámszerűségének rugalmasabb kezelését a normalizált időtartam függvényében. A modellt a beszélői csoport parametrikus faktor változó fix hatás mellett minden tónus esetében random simítással (random smooth, bs = "fs") bővítettem. Ez két különböző random smooth funkciót jelentett: egyet a beszélőkre illesztve, egyet pedig a külön álló token-megvalósulások szerint. Az autokorreláció ellenőrzése (acf.resid()) után a modellek által becsült f0-görbék minden esetben 95%-os konfidencia-intervallummal ábrázoltam az itsadug R csomaggal (van Rij et al., 2020).

3. Eredmények

Az eredmények bemutatását a lexikai tónusok időtartamára vonatkozó vizsgálattal kezdem, majd ezt követően mutatom be az f0-kontúrokat középpontba helyező dinamikus elemzést.

3.1. A kínai lexikai tónusok időtartamának összehasonlítása a három beszélői csoport között

A négy kínai lexikai tónus időtartamát illetően abszolút és relatív időtartam-értékeket mutatok be. Az abszolút időtartam-értékeket pusztán csak referenciaként és kvalitatív szempontból mutatom be, majd ezt követően térek rá a relatív időtartamokra, amelyeket már statisztikailag megerősített vizsgálat is kísér.

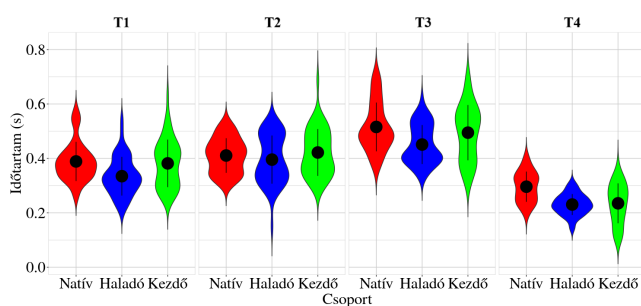
Az abszolút időtartamértékek tekintetében az 5. ábrán azt láthatjuk, hogy az adatok minden beszélői csoport esetében jelentősen szóródnak, azonban megfi-

gyelhetünk egy általános tendenciát a csoportok között: a lexikai tónusok között a T3 rendelkezik minden csoport esetében a leghosszabb, míg a T4 rendelkezik a legrövidebb időtartamokkal (2. táblázat). A T1 és a T2 hasonló átlagértékekkel, de köztes időtartamokkal realizálódik a T3-ra és a T4-re jellemző átlagértékekhez képest, habár a T2 minden csoport ejtésében a T1-hez képest árnyalatnyival hosszabb időtartammal realizálódott (5. ábra). Ha a négy tónus ejtését globálisan tekintjük a csoportok között, akkor az a mintázat látszik kirajzolódni, hogy – ugyan csak tendenciák szintjén – de a csoportok között minden tónus esetében a haladó nyelvtanulók rendelkeznek a legalacsonyabb átlagértékkel, míg a kezdők ehhez képest megközelítik a natív beszélők időtartamértékeit (5. ábra).

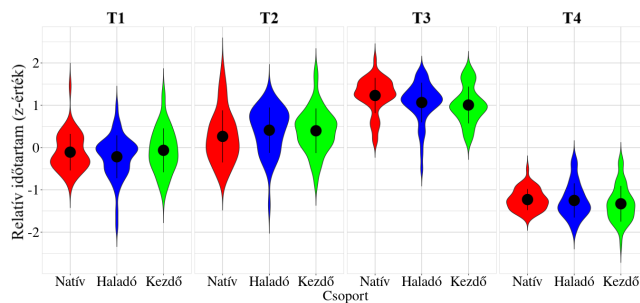
A relatív időtartamok (6. ábra) esetében az előbb leírt időtartam-különbségek a négy lexikai tónus között nyilvánvalóbbá válik, azonban a csoportok átlagában megfigyelhető mintázatok is eltérően alakulnak (2. táblázat). Mindhárom beszélői csoport esetében a T1 és a T2 ejtésének időtartama közelítette meg a legjobban a beszélőkre külön-külön számolt átlagértékét ($z = 0$) a négy lexikai tónusra vonatkozóan. Míg a három beszélői csoport a T1-et az átlagértékhez képest egy árnyalatnyival rövidebben, addig a T2-t éppen fél szórással hosszabb időtartammal ejtették. A lexikai tónusok között a T3 volt az a tónus, amelynek ejtése az átlagos ejtési időtartamhoz képest 1 szórással hosszabb átlagértékkel valósult meg, míg a T4 esetében az átlagos ejtési időtartamhoz képest 1,5 szórással rövidebb időtartam figyelhető meg. A csoportok között a statisztikai próbák nem mutattak szignifikáns eltérést egyik tónus időtartamértékeiben sem. Ez azt jelenti, hogy a két kínaiul tanuló magyar anyanyelvű csoport mind a négy lexikai tónust a kínai anyanyelvűekkel megegyező időtartammal ejtette. Tendenciaszintű különbségek azonban megfigyelhetők a csoportok között: a T3 és a T4 esetében a kezdő nyelvtanulók produkálták a csoportok közötti legrövidebb relatív időtartamértékeket, míg a T1 és a T2 esetében ezzel ellenkezőleg éppen a kezdők produkálták a csoportok közötti legmagasabb relatív értékeket.

2. táblázat. A négy kínai lexikai tónus abszolút időtartamának átlaga (bal táblázat), illetve az abszolút átlagidőtartamok beszélőnként kiszámított, z-érték mentén normalizált átlagértéke.

abszolút időtartam (s)	natív	haladó	kezdő	relatív időtartam	natív	haladó	kezdő
T1	0,39	0,33	0,38	T1	-0,10	-0,22	-0,065
T2	0,41	0,40	0,42	T2	0,26	0,41	0,39
T3	0,52	0,45	0,49	T3	1,23	1,07	1,01
T4	0,30	0,23	0,23	T4	-1,23	-1,25	-1,33



5. ábra. A négy kínai tónus abszolút időtartama a három beszélői csoport ejtésében (ahol a fekete pont az átlagértéket, míg a fekete vonal a szórást jeleníti meg).



6. ábra. A négy kínai tónus relatív időtartama a három beszélői csoport ejtésében, vagyis az időtartam-értékek a beszélők szerint a z-érték mentén normalizáltak (ahol a fekete pont az átlagértéket, míg a fekete vonal a szórást jeleníti meg).

3.2. Az f0-kontúrok dinamikus összehasonlítása a három beszélői csoport között

3.2.1. A magas szinttartó T1 tónus

A három beszélői csoport T1 f0-kontúrait összehasonlító GAMM-modell a következő eredményt hozta: A T1 modell parametrikus eredményei alapján azt mondhatjuk, hogy az egész tónuskontúr f0-értékeinek átlagát tekintve mindkét nyelvtanuló csoport ejtése a natív mintával megegyező f0-tartományban valósult meg (3. táblázat). A görbék alakját véve középpontba, a smooth együtthatók alapján mindkét nyelvtanulói csoport szignifikánsan eltért a natív kontúr alakjától (haladók: $p < 0,05$; kezdők: $p < 0,001$, 3. táblázat). Ha a T1-kontúrok alak megvalósulását vetjük össze az 7. ábrán, akkor az látható, hogy mindkét nyelvtanulói csoport a natív ejtésre jellemző, relatíve magas f0-tartományban megvalósuló stagnáló mintázatot produkál. A haladó nyelvtanulók ejtését a natív f0-kontúrral közel teljesen átfedő egyenes jellemzi, ahol a két f0-lefutás becsült különbsége kisebb, mint 1 félhang (7. ábra). A kezdő nyelvtanulók esetében az f0-görbe a natív mintához hasonlóan szép egyenes, azonban a natív kontúrhoz képest majdnem 1 félhanggal magasabb értékekkel realizálódik (3. táblázat, 7. ábra).

3.2.2. Az emelkedő T2 tónus

A T2-re illesztett modell eredményei alapján azt láthatjuk, hogy a parametrikus eredmények szerint a görbe átlagos f0-értéke mentén egyik csoport sem tért el a natív f0-kontúrtól (3. táblázat). Továbbá a haladó nyelvtanulók a görbe alakjában, vagyis a smooth funkciók tekintetében sem tértek el a kínai anyanyelvű beszélőktől, tehát e tónust mind az f0-tartomány, mind az f0-lefutás tekintetében az anyanyelvi kínai ejtéssel megegyezően ejtették, még akkor is, ha maga a görbe a natív ejtéshez képest egész lefutásában a natív ejtésnél 1,5 félhanggal alacsonyabb f0-tartományban realizálódott. A kezdő nyelvtanulók f0-kontúrája azonban a natív ejtésnél árnyalatnyival (kevesebb, mint 1 félhanggal) magasabb f0-értékekkel indult a normalizált időtartam elején, míg a natív mintához képest közel 2 félhanggal alacsonyabb f0-értékekkel zárult (8. ábra), amely eltérések a natív mintánál több, mint 1,2 félhanggal alacsonyabb,

az egész kontúrra számított becült átlagértéket eredményezett (3. táblázat). Továbbá azt is meg kell jegyezni, hogy a kontúr homorúbb fázisa a natív és haladó kínaiul tanuló beszélők ejtéséhez képest a normalizált időtartamon belül később következett be (7. ábra). Mindezen alaki tulajdonságok eredményezheték a szignifikáns eltérést a smooth funkciók tekintetében ($p < 0,01$) a kezdő nyelvtanulók ejtésében. Ha az emelkedés mértékét félhangokban fejezzük ki a csoportok között a görbék alapján, akkor azt mondhatjuk, hogy a natív és a haladó nyelvtanuló csoport esetében 6-6 félhangnak, tehát egy bővített kvartnak felel meg, míg a kezdők ejtésében az emelkedés megközelítőleg 5 félhangot, vagyis egy tiszta kvartot jelentett.

3.2.3. A mély ereszkedő-emelkedő T3 tónus

A T3 ejtését véve középpontba (7. ábra) egy relatíve alacsony f0-tartományban megvalósuló homorú görbét figyelhetünk meg mindhárom beszélői csoport ejtésében. A GAMM eredményei szerint az f0-görbe becült átlagát tekintve a két nyelvtanulói csoport egyike sem tért el a natív ejtéstől, azonban – ahogy a parametrikus becült átlagértékek is megjelenítik – a csoportok között a kezdő nyelvtanulók produkálták a legmagasabb f0-tartományban a dallamkontúrt, amely görbe kezdeti és záró fázisa a natív ejtést megközelítő becült értékekkel realizálódott. Azonban a kezdő nyelvtanulók ejtésében a kontúrt a natív ejtéshez képest kevésbé homorú mintázat jellemezte ($p < 0,05$), vagyis összességében a natív ejtésnél magasabb f0-értékek kísérték. A haladó nyelvtanulók ejtésében a kezdők mintázatánál is lapultabb f0-kontúrt figyelhetünk meg ($p < 0,001$), amely görbe kezdeti és záró f0-értékei is a másik két csoporthoz képest alacsonyabb f0-értékekkel jelentek meg. Továbbá az f0-görbe alakját megfigyelve azt láthatjuk, hogy az adaptív smooth funkció ellenére a haladó nyelvtanulók kontúrja relatíve hullámzó. Ez a girbe-gurba forma elsősorban a normalizált időtartam középső fázisában jelentkezik, ahol a másik két csoport kontúrjának legalacsonyabb f0-értékeit figyelhetjük meg és ezzel párhuzamosan az irreguláris zöngéképzés okozta zajt is feltételezhetjük az adatokban. Az ereszkedő-emelkedő fázis nagyságát helyezve középpontba, a natív és kezdő nyelvtanulói ejtésben

megközelítőleg szimmetrikus görbéket figyelhetünk meg, amelyek a normalizált időtartam felénél érik el minimum pozíciójukat. A natív ereszkedés-emelkedés megközelítőleg 9 félhangot jelentett, amely egy nagy szextnek feleltethető meg, míg a kezdők ejtésében ez inkább csak 4 félhang, vagyis egy nagy terc hangmagasságváltozást jelentett. A haladó nyelvtanulók esetében az f_0 -változás leírása a görbe egyenletlenségeiből fakadóan egy fokkal nehezebb: a kezdeti ereszkedés 4 félhangra, vagyis egy nagy tercre tehető, ami nagyságrendileg megegyezik a kezdők értékeivel, illetve ha ettől a ponttól számítjuk az emelkedést is, akkor ugyanúgy 4 félhangra tehető. E lexikai tónus esetében figyelhető meg az egyetlen – GAMM szerint relevánsnak ítélt – markáns eltérés a natív és a kezdő kontúrok között: a normalizált időtartam 61,6% és 65,6% százaléka között a két f_0 -görbe közötti eltérés kb. 2 félhangnyira volt becsülhető (8. ábra). Habár ebben az esetben azt láthatjuk, hogy a modell predikciója szerint a fentebb említett fázis már szignifikánsan eltér, azonban Sós-kuthy (2021) tanulmányára és személyes tanácsára hivatkozva ezen 8. ábrán piros sávval megjelenített eltérést kezeljük kitüntetett szereppel és bármely más 8. ábrán megjelenített eltéréstől eltérően. E döntés a szignifikánsan eltérő fázis nagyságára és a pozíciójára is alapozható: a tónuskontúrok mindössze 4%-át érintő eltérés aligha tekinthető megbízhatóan széles intervallumnak, amelyre a szignifikancia-jelölésében hagyatkozni lehetne. A prediktált eltérés inkább pontszerű különbségeket jelez, amelyek éppen abban a fázisban mutatkoznak, ahol az irreguláris zöngé legnagyobbnal valószínűséggel feltételezhető, aminek révén a különbség relevanciája megkérdőjelezhető.

3.2.4. Az ereszkedő T4 tónus

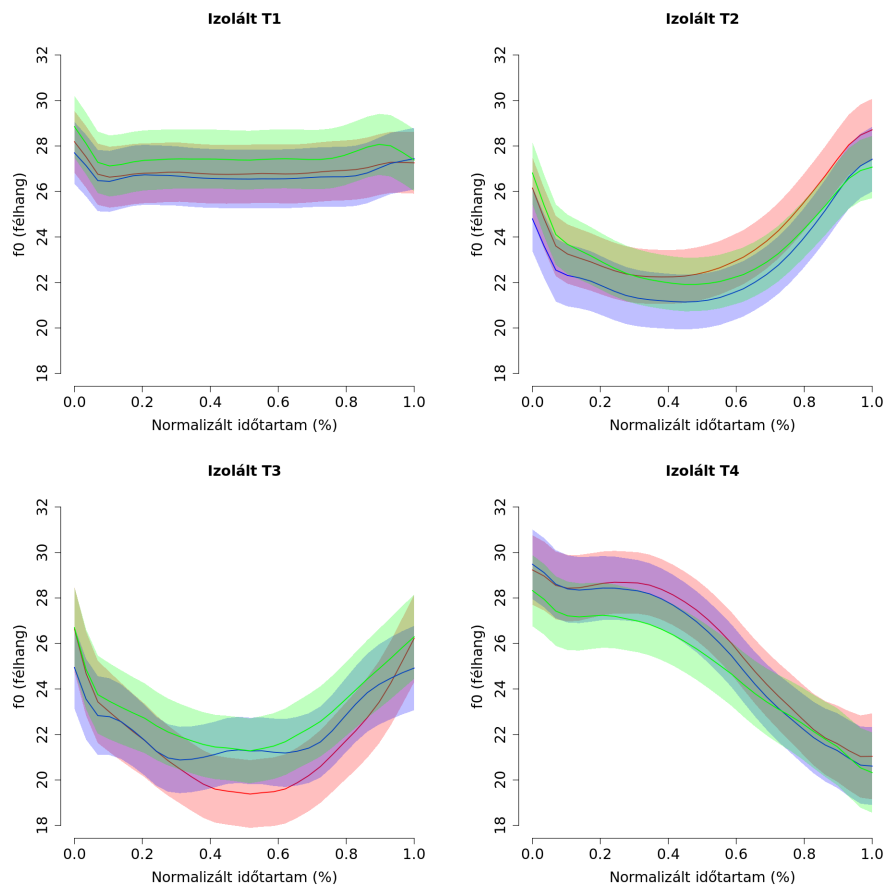
A T4 lexikai tónus esetében mindhárom beszélői csoport esetében nyilvánvaló az f_0 -értékek ereszkedése a vizsgált szótag normalizált időtartamán belül. E lexikai tónus esetében a statisztikai próba eredményei szerint egyik nyelvtanuló csoport sem tért el a natív ejtéstől sem a kontúr f_0 -tartományában, sem a görbe alakjában (3. táblázat). Azonban a 7. ábrán azt láthatjuk, hogy a kínai anyanyelvi beszélők és a haladó nyelvtanulók f_0 -görbéi a normalizált időtartam legelején teljesen átfednek, majd megközelítőleg a normalizált időtartam

20%-ától a haladók egy árnyalatnyival alacsonyabb f0-átlagértékeket és ennek következtében egy fokkal kevésbé domború f0-görbét produkálnak. Ehhez képest a kezdő nyelvtanulók f0-kontúrja a másik két csoporthoz viszonyítva kb. 1 félhanggal alacsonyabb f0-tartományból indul ereszkedésnek, azonban a kontúr záró fázisára az átlagértékek megközelítik a natív, illetve haladó csoport ejtési mintázatát (7. ábra, 8. ábra). Az ereszkedés mértéke a natív és haladó nyelvtanuló beszélőknél 8 félhangra (vagyis egy bővített kvintre/kis szextre) volt tehető, míg a kezdő nyelvtanulók esetében ez inkább 7 félhangnak (egy tiszta kvintnek) felelt meg.

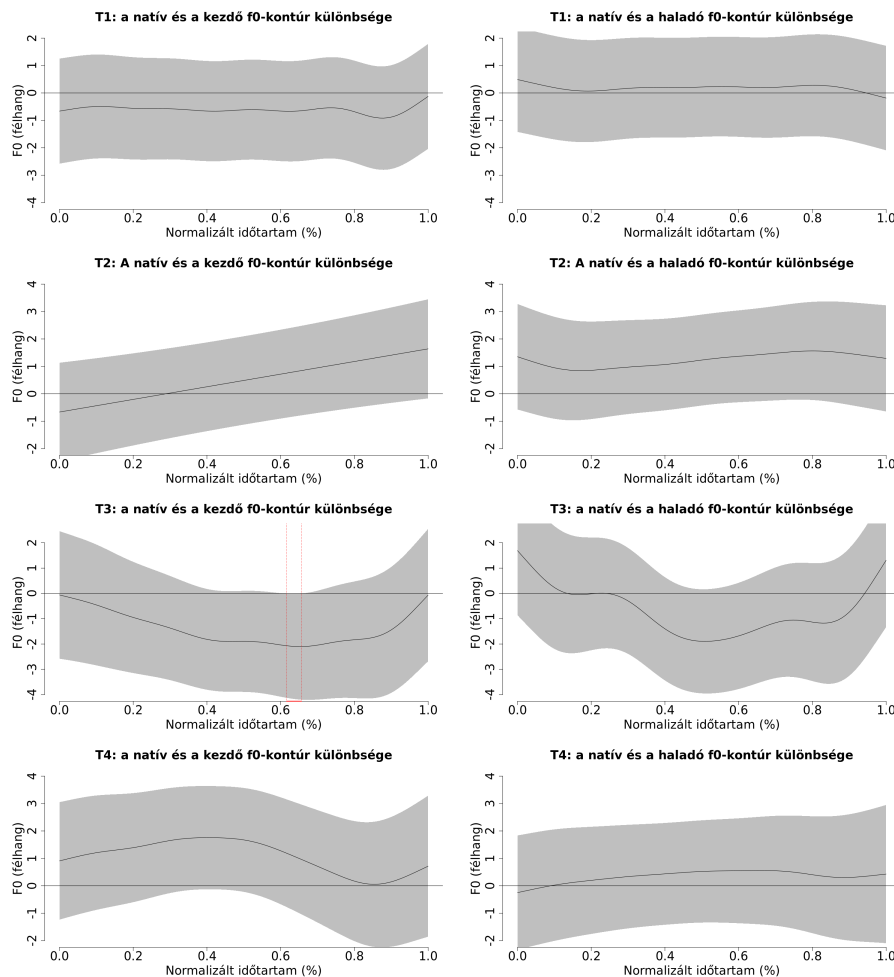
Továbbá, ami összességében mind a négy lexikai tónus statisztikai eredményeivel kapcsolatban elmondható, hogy tónuskontúrtól függetlenül a beszélőkre és token-megvalósulásokra illesztett random smooth funkciók egytől-egyig szignifikáns eredményt hoztak (3. táblázat, $p < 0,001$), amely eredmények jelentős variabilitásra utalhatnak mind a beszélők mentén, mind pedig a külön-külön megvalósuló tónuskontúrok lefutását tekintve.

3. táblázat. A GAMM-ok parametrikus és smooth együtthatói (ahol a **beszélő** a beszélőnként, illetve a **traj.** a tokenenként illesztett random smooth-ot jeleníti meg és az n.s. jelölés a nem szignifikáns eredményt jelöli).

T1 ($R^2 = 98,6\%$)				T2 ($R^2 = 99,0\%$)			
parametrikus együtthatók				parametrikus együtthatók			
	becsült átlag	t	$Pr(> t)$		becsült átlag	t	$Pr(> t)$
natív	26,9	41,5	<0,001	natív	24,0	41,2	<0,001
haladó	26,7	-0,2	n.s.	haladó	22,8	-1,5	n.s.
kezdő	27,5	0,7	n.s.	kezdő	23,5	-0,6	n.s.
smooth együtthatók				smooth együtthatók			
	edf	F	p		edf	F	p
natív	29,0	78,9	<0,001	natív	29,1	306,6	<0,001
haladó	5,6	2,2	<0,05	haladó	5,2	1,3	n.s.
kezdő	7,8	7,2	<0,001	kezdő	6,6	3,0	<0,01
beszélő	183,4	17,1	<0,001	beszélő	165,6	5,7	<0,001
traj.	508,6	20,1	<0,001	traj.	977,7	30,8	<0,001
T3 ($R^2 = 97,7\%$)				T4 ($R^2 = 99,5\%$)			
parametrikus együtthatók				parametrikus együtthatók			
	becsült átlag	t	$Pr(> t)$		becsült átlag	t	$Pr(> t)$
natív	21,9	33,6	<0,001	natív	26,8	40,8	<0,001
haladó	22,4	0,5	n.s.	haladó	26,5	-0,4	n.s.
kezdő	23,4	1,4	n.s.	kezdő	25,7	-1,3	n.s.
smooth együtthatók				smooth együtthatók			
	edf	F	p		edf	F	p
natív	23,1	64,3	<0,001	natív	26,0	18,9	<0,001
haladó	7,4	13,1	<0,001	haladó	3,7	0,4	n.s.
kezdő	5,6	2,03	<0,05	kezdő	5,8	1,8	n.s.
beszélő	155,4	7,1	<0,001	beszélő	125,3	3,5	<0,001
traj.	852,1	16,1	<0,001	traj.	1064,0	71,6	<0,001



7. ábra. A négy kínai lexikai tónus (T1 = balra felül, T2 = jobbra felül, T3 = balra alul, T4 = jobbra alul) GAMM-modellek eredményeként becsült f_0 -kontúrja, ahol a piros szín a natív kínai, a kék szín a haladó nyelvtanulók, míg a zöld szín a kezdő nyelvtanulók csoportjának produkcióját jelöli.



8. ábra. A két vizsgált nyelvtanulói csoport (haladók, kezdők) natív f₀-kontúroktól szignifikánsan eltérő intervallumai a normalizált időtartam függvényében a négy kínai lexikai tónus ejtésében, ahol a bal oszlopban a kezdők, a jobb oszlopban a haladók differenciagörbéi láthatók.

4. Következtetések

A tanulmány az izolált ejtésű mandarin kínai lexikai tónusok produkciójának átfogó akusztikai fonetikai vizsgálatát tűzte ki célul a kínaiul tanuló magyar anyanyelvű beszélők ejtésében. A tanulmányban a négy kínai lexikai tónus produkcióját két akusztikai szempont szerint, vagyis a dallamívük menete, valamint

a megvalósulásuk időtartama szerint hasonlítottam össze két, eltérő mértékű L2 nyelvi tapasztalattal rendelkező kínaiul tanuló csoport ejtésében. A szótag egészének az időtartamát lineáris kevert modellekkel, az f0-görbék lefutását GAMM-okkal végeztem. Az első kutatási kérdés a lexikai tónusok időtartamára vonatkozott, vagyis hogy a kínaiul tanuló magyar anyanyelvűek a kínai beszélőkkel megegyező időtartammal ejtik-e a négy mandarin lexikai tónust. Illetve, ha a nyelvtanulók eltérnek a natív mintázatoktól, akkor az eltérésben megfigyelhető-e a célnyelvi tapasztalat mennyiségének hatása. A statisztikai elemzést a lexikai tónusok relatív időtartamán végeztem, mert az időtartam-értékek beszélők mentén való normalizációja visszaszorította a kísérleti személyek beszédtempójában megfigyelhető eltéréseket is. A lexikai tónusok időtartamában célnyelvi tapasztalattól függetlenül egyik nyelvtanulói csoport sem tért el a kínai natív ejtéstől. Ez azt jelenti, hogy a magyar nyelvtanulók képesek a natív mintának megfelelően produkálni a mandarin lexikai tónusok temporális sajátosságait, és ezen jelenségre nincs hatással a célnyelvi tapasztalat mennyisége. A relatív átlagos időtartamok kvalitatív eredményei abba az irányba mutatnak, hogy a rendszer szintjén mindkét nyelvtanuló csoport megközelítőleg fele olyan hosszan ejti a T4-et, mint az összehasonlítás referenciájaként szolgáló T3-at, tehát a natív beszélőkhöz hasonlóan különítik el e két tónust (de ezen állítást a jelen kísérlet empirikus úton nem erősítette meg). Tehát az izolált ejtésű kínai lexikai tónusokat mindkét nyelvtanuló csoport a natív időtartam-mintának megfelelően ejtette, azonban ezen tónusok időtartam alapján megvalósuló akusztikai elkülönítése nyitott kérdés marad. A négy kínai lexikai tónus közötti időtartam-kontraszt részletes, kvantitatív vizsgálatát egy előkészületben lévő tanulmányomban mutatom be. Mindenesetre egy rövid gondolat erejéig arra is érdemes kitérni, hogy mi teszi lehetővé azt, hogy a nyelvtanulók a natív mintázatnak megfelelő időtartamokkal ejtik a kínai tónusokat. Erre a jelenségre például magyarázattal szolgálhat az anyanyelvi mintázatok megfigyelése: habár a szerző nincs tudomással olyan összefoglaló tanulmányról, ahol a magyar egy szótagú megnyilatkozások időtartamát vetették volna össze akár percepció, akár produkció tekintetben, azonban például Juhász (2023) eredményei szignifikáns eltérést mutattak a fel-

szólító és a kijelentő monoszillabikus egységek időtartamában. Ezen eredmények – illetve emellett mindennapi tapasztalataink – alapján feltételezhető, hogy az f0-kontúr megvalósulása mellett a közlés időtartama is hozzájárulhat a jelentés kialakításához, és adott esetben feltételezhetően kontrasztív akusztikai tényező is lehet a magyar nyelvhasználatban. Ebből fakadóan számíthatunk arra, hogy a magyar anyanyelvűek érzékenyek ezen akusztikai jegy eltéréseire, ami elősegíthette a natív ejtésnek megfelelő időtartam-produkciót a kínai lexikai tónusok esetében. A jövőben mindenképpen érdemes lenne azon kutatási kérdés vizsgálata, hogy a magyar anyanyelvű beszélők észlelésében milyen hatással rendelkezik, és ejtésükben hogyan alakul a monoszillabikus megnyilatkozások időtartama.

A lexikai tónusok dallamívének megvalósulásával kapcsolatban kérdésként merült fel, hogy a magyar beszélők a kínai natív dallamkontúrokhöz hasonlóan produkálnak-e vagy eltérnek a kínai dallamkontúroktól. Illetve eltérés esetében e akusztikai tulajdonság vizsgálatában is felvethető az a kérdés, hogy az eredményekben a nyelvi tapasztalat hatása is megfigyelhető-e, vagyis a haladók jobban megközelítik-e a natív mintázatokat, mint a kevesebb L2-tapasztalattal rendelkező kezdő nyelvtanulók.

A magas szinttartó T1 esetében azt láthattuk, hogy a dallammenet pozicionálása, pontosabban a kontúr f0-tartománybeli magassága nem okozott problémát a nyelvtanulóknak, hiszen mindkét csoport átlagosan ugyanolyan f0-értékekkel ejtette ezt a statikus magas lexikai tónust. Az f0-görbe alakját tekintve a haladók alapvetően nem tértek el a natív kínai ejtéstől, a két csoport között megfigyelhető eltérés nem érte el az 1 félhangot sem. A kezdő nyelvtanulók ejtésében a natív mintáznál árnyalatnyilag (1 félhanggal) magasabb f0-értékeket láthattunk a teljes tónuskontúr megvalósulását illetően, azonban a két csoport f0-egyenes a normalizált időtartam egészében átfedett egymással. Ez azt jelenti, hogy mindkét nyelvtanulói csoport a natív ejtéssel megegyezően ejtette a T1-es tónust. A nyelvi tapasztalatra vonatkozó kutatási kérdés e tónus esetében megerősíteni látszik, hogy a haladók kisebb mértékben tértek el a natív mintától, mint a kevesebb nyelvi tapasztalattal rendelkező kezdő nyelvtanulók. A kezdő nyelvtanulók kínai anyanyelvi ejtéshez képest „túllőtt” magas akusztikai

célja (target overshoot) fakadhat a beszédfeladat minőségéből és a megfigyelés tényéből is. Empirikus kísérletek esetében a Labov-féle megfigyelési paradoxonra (1972) alapozva feltételezhetjük, hogy a beszélők máshogy beszélnek spontán helyzetben, és máshogy akkor, amikor éppen figyelik őket. A jelen kísérlet a laboratóriumi akusztikai kísérletek legkontrolláltabb formája, ahol a beszélő figyelve egyetlen szótag kiejtésére összpontosul, ráadásul idegen nyelven, úgy, hogy a kísérlet vezetője mellette ül és folyamatosan figyeli. Ezen körülmények sok beszélőben kelhetnek bizonyos mértékű szorongást a potenciális értékelés fényében (Horwitz et al., 1986), még akkor is, ha a kísérletvezetőnek ez nem célja és próbál a lehető legkellemesebb és legnyugodtabb környezetet teremteni. Emellett egy pillanatra visszatérve a laboratóriumi kísérlet feladattípusára, az izolált ejtésű felolvasási kísérlet esetében – mint korábban már említettem – hiperartikulált ejtésre számíthatunk, vagyis a prominens akusztikai jegyek túlzó megvalósítására (vö. Scarborough & Zellou, 2013). A kezdő nyelvtanulók esetében a T1-re vonatkoztatva ez feltételezhetően azt jelentette, hogy a T1 legprominensebb, magas f0-tartományban megvalósuló akusztikai tulajdonságait túlartikulálták és ezért valósult meg a kontúr a natív ejtésnél is magasabb f0-értékekkel.

Az emelkedő T2 tónus elsajátítása is sikeres volt mindkét nyelvtanuló csoport esetében: a haladó nyelvtanulók minden tekintetben a natív mintának megfelelően ejtették e kínai lexikai tónust, mégis a haladók teljes f0-kontúrja a natív mintánál kb. 1,5 félhanggal alacsonyabban realizálódott. A natív ejtésnél alacsonyabb f0-értékek, de a megegyező alakú f0-görbe megvalósulása nem az anyanyelvi magyar mintázatok eredményének tűnik, hiszen ha a magyar egy szótagú eldöntendő kérdő dallamot vesszük alapul, a tónus-görbe jelentős hányadát a kínai tónusnál magasabb f0-értékek jellemezték, szemben a haladók ejtésében megvalósuló alacsony f0-értékekkel. Ebben az esetben talán inkább egy másik irányban érdemes a mintázatok mögött megbúvó motivációt kutatni. Habár a jelen tanulmánynak nem célja a tónusok közötti kontraszt vizsgálata, azonban a bemutatott eredmények egyértelműen felvetik ennek a kérdésnek a relevanciáját. Mind a haladó, mind a kezdő kínaiul tanulók a T2-t a natív

kínai mintánál alacsonyabb f_0 -értékekkel, míg a T3-at pedig a natív mintánál magasabb f_0 -értékekkel ejtették. Ezek az akusztikai tulajdonságok felvetik a kérdést – amely kérdés vizsgálatára egy következő tanulmányban kerítünk sort – hogy vajon nem a T2 és a T3 ejtésbeli megkülönböztetésének hiánya indukálja-e ezeket a nyelvtanulói mintázatokat. Hiszen ahogy Hao (2012) is összefoglalta: a mindkét tónusra jellemző ereszkedő-emelkedő mintázat könnyen eredményezheti azt, hogy e két tónus nem különül el megfelelően produkciósan (és talán percepciósan sem). A T2 és a T3 tónus esetében megfigyelhető mintázatok is összhangban állnak Hao (2012) eredményeivel, hiszen e két tónus esetében a natív ejtéstől való eltérés inkább 2 félhangra volt tehető, míg például a T1 esetében ez nagyjából maximum 1 félhang volt, vagyis a T2 és a T3 esetében jelentősebb eltérések figyelhetők meg a magas szinttartó T1-hez képest.

Az ereszkedő-emelkedő T3 esetében azt láthattuk, hogy mindkét nyelvtanuló csoport görbéjének f_0 -terjedelme a kínai anyanyelvű ejtéshez képest kompresszáva, és kevésbé homorúan valósult meg, amit a natív ejtéshez képest magasabb minimális f_0 -érték idézett elő. Továbbá, a T3 esetében azt láthattuk, hogy a natív ejtés f_0 -görbéje a relatíve alacsony f_0 -értékek ellenére is egyenesen homorú mintázatot mutatott, szemben például a haladó nyelvtanulók tónuskontúrjával. Ezen ejtési tulajdonság például annak a képzési jegynek is tulajdonítható, hogy míg a natív beszélők esetében az irreguláris zöngé produkciója jól kontrollált, addig a haladó nyelvtanulók esetében felmerülhet, hogy az irreguláris zöngéképzés révén létrejövő f_0 -értékek sokkal kevésbé kontrolláltak, vagyis zajosabbak az adatok. A kezdő nyelvtanulók esetében az f_0 -kontúr magasabb f_0 -tartományban valósult meg, így feltételezhetően az irreguláris zöngé indukálta hatás is visszafogottabb a haladó nyelvtanulók ejtéséhez képest. Az irreguláris zöngé akusztikai elemzése egy következő elemzés kutatási kérdéseként jelenhet meg, vizsgálva az irreguláris zöngé intervallumának pozícióját, hosszát és az f_0 -értékek minőségét a kínaiul tanuló magyar anyanyelvűek ejtésében. Mindent összevetve azonban a haladó nyelvtanulók – még ha az f_0 -görbe alakja szempontjából nem is, de – jobban megközelítették a natív ejtést, mint a kezdő nyelvtanulók, így tehát e tekintetben a Flege- és Bohn-féle predikció

e tónus esetében is megerősítést nyert. Ha a T3-hoz kapcsolódó eredményeket a szakirodalmi forrásokkal hasonlítjuk össze, akkor a Chao (1948/1963) által lejegyzett bővített kvart (tehát 6 félhang) helyett 9 félhangnyi f₀-terjedelmet figyelhetünk meg. Ezzel szemben például Chun és munkatársai (2015: 3) mérésében a megközelítőleg 160 Hz-es minimumértéket és a 280 Hz-es maximumértéket 50 Hz-es referenciaértékkel félhangokká konvertálva 20 és 29,8 félhangot kapunk, amely f₀-terjedelem megegyezik a jelen kísérletben bemutatott eredményekkel. Ezen összehasonlítás alapján felmerül a kérdés, hogy vajon a Chao-féle eredmények (1948/1963), illusztrációk relatíve régi, illetve saját megfigyelésen nyugvó leírásainak használata pozitívan járulhat-e hozzá a kínaiul tanulók lexikaitónus-produkciójához. E kérdést a T4 nyelvtanulói produkciója is felveti.

Az ereszkedő T4 esetében is megerősíthető a Flege- és Bohn-féle (2021) predikció arra vonatkozóan, hogy a haladó nyelvtanulók jobban megközelítik az anyanyelvi mintázatot a kevesebb L2-tapasztalattal rendelkező kezdő nyelvtanulókhoz képest, hiszen a kezdők alacsonyabb f₀-tartományból ereszkedve kisebb terjedelmű kontúrt produkáltak a natív beszélőkhöz képest. Ezeket az eredményeket potenciálisan magyarázhatjuk az anyanyelvi magyar felszólító dallamkontúr hatásával, hiszen az alacsonyabb f₀-terjedelemről indul ki a kínai T4-hez képest, azonban e lehetséges hatás szerepe elég visszafogott, hiszen mindössze egy félhang különbséget eredményez. Ezen ejtési tulajdonságot akár a lexikai tónusok számkódjainak hatására is visszavezethetjük: a bevezetés részben említettem, hogy ellentétben azzal, hogy a T4 a T1-hez képest magasabb f₀-tartományból indul, a kínai nyelvtanításban mindkét tónus 5-ös hangmagassági szinttel van jellemezve az 5 fokozatú skálán, ami negatív hatással lehet a nyelvtanulók produkciójára. Ebben az esetben elképzelhető, hogy pont ezt a hatást figyelhetjük meg: mind a T1, mind a T4 megközelítőleg 28,5 félhang magasságból indul, tehát a két tónus kezdeti akusztikai célja a kezdő nyelvtanuló csoportok ejtésében megegyezik. E megfigyelés megerősítéséhez azonban a már korábban említett tónus-kontrasztok vizsgálata szükséges. Ugyanakkor visszatérve a lexikai tónusok elsajátításának relatív nehézségére a nyelvtanulók körében, a T4 esetében a T1-hez hasonló mértékű eltéréseket figyelhetünk meg,

amely eltérés a haladók esetében kisebb volt mint 1 félhang, és a kezdő nyelvtanulók esetében sem érte el a különbség a 2 félhangot (szemben a T2 és a T3 esetében tapasztaltakkal, ahol akár 2 félhangot meghaladó eltérések is megfigyelhetők voltak). Mindezen eredmények összhangban látszanak lenni Hao (2012) eredményeivel, amelyek szerint a T1 és a T4 elsajátítása sikeresebb a T2-höz és a T3-hoz képest, hiszen utóbbi esetekben jelentősebb eltéréseket figyeltünk meg a natív mintázathoz képest mint a T1 és a T4 esetében. A fentebbiek alapján összegezve a célnyelvi tapasztalat jelentőségére vonatkozó kutatási kérdést, azt mondhatjuk el, hogy minden tónus esetében megfigyelhető volt az a mintázat, hogy a haladók kisebb eltérésekkel, vagyis jobban megközelítik a natív mintázatot, mint a kevesebb célnyelvi tapasztalattal rendelkező kezdő nyelvtanulók. Tehát összességében a kutatás eredményei az f0-kontúrok esetében megerősítik Flege és Bohn (2021) predikcióját, miszerint a nyelvi tapasztalat növekedésével a nyelvtanulók célnyelvi teljesítménye jobban megközelíti a natív célnyelvi mintázatokat. Ezzel szemben az időtartam tekintetében nem látjuk ezen eltéréseket a nyelvtanulói csoportok között a lexikai tónusok produkciójában. E jelenség alapján az feltételezhető, hogy nem meglepő módon a temporális sajátosságok, pontosabban a tónus időtartamának elsajátítása könnyebb és hatékonyabban megy végbe a nyelvtanulók számára, szemben a hangszalagműködés finommotorikán nyugvó beállításával és szabályozásával. Habár az időtartam is skaláris tulajdonságként számszerűsíthető, azonban ez esetben kizárólag a zöngképzés időtartamát (kezdetét/fenntartását/megszűnését) kell szabályozni, míg a dallamkontúrok esetében egy sokkal bonyolultabb akusztikai célt kell megtalálni, amelyhez a jelen kísérletben – az izolált ejtésű megnyilatkozásokból fakadóan – semmilyen kontextuális segítség nem állt rendelkezésre. Ezek az eredmények közvetetten összhangban állnak Flege és Bohn (2021) Revised speech learning modeljében leírtakkal, miszerint a komplexebb mintázatok (e modell esetében ez elsősorban beszédhangokra értendő) elsajátítása nehezebben megy végbe az egyszerűbb mintázatokhoz képest. A tanulmányban feltárt akusztikai eltérések jelentősége a jövőben percepciók tesztek segítségével vizsgálható. Továbbá kérdésként merül fel, hogy a négy kínai tónus – a dallamok rendszere szintjén

– mennyire különül el a kínaiul tanuló magyar anyanyelvűek ejtésében, amely kérdésre egy előkészületben lévő tanulmányban keresem a választ. A kísérlet eredményei hozzájárulnak a kínai tónusok elsajátításának és ejtési problémáinak mélyebb megértéséhez és elősegíthetik a kínai mint idegen nyelv oktatását.

Köszönetnyilvánítás

Hálával tartozom Sós-kuthy Mártonnak, akinek a rendkívül részletes és segítőkész bírálatát olvasva sokat tanulhattam többek között a GAMM-ok használatáról és működéséről, amely ismereteket a tanulmány statisztikai módszertanában is alkalmaztam. Általánosságban köszönettel tartozom a tanulmány bírálóinak, akik javaslataikkal és kérdéseikkel segítettek érthetőbbé és fókuszáltabbá tenni a tanulmányt. A Kulturális és Innovációs Minisztérium EKÖP-24 kódszámú Egyetemi Kiválósági Ösztöndíj Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.

Hivatkozások

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. doi:10.18637/jss.v067.i01.
- Boersma, P., & Weenink, D. (2022). Praat: doing phonetics by computer. [Computer program]. 6.3.03-as verzió. (letöltés ideje: 2022. december 17.).
- Bolla, K. (1995). *Magyar Fonetikai Atlasz*. Budapest: Nemzeti Tankönyvkiadó.
- Chao, Y. (1948/1963). *Mandarin Primer*. Cambridge: Harvard University Press. URL: <https://doi.org/10.4159/harvard.9780674732889>. doi:10.4159/harvard.9780674732889.
- Chun, D., Jiang, Y., Meyr, J., & Yang, R. (2015). Acquisition of l2 mandarin chinese tones with learner-created tone visualizations. *Journal of Second Language Pronunciation*, 1, 86–114. doi:10.1075/jslp.1.1.04chu.

- Dallaston, K., & Docherty, G. (2019). Estimating the prevalence of creaky voice: a fundamental frequency-based approach. In *International Congress of Phonetic Sciences (ICPhS)* (pp. 532–536). Melbourne, Australia.
- Dorreen, K. (2017). *Fundamental frequency distributions of bilingual speakers in forensic speaker comparison. Mesterszakos szakdolgozat*. Christchurch: The University of Canterbury.
- Flege, J. (1995). Second language speech learning: theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 233–277). Timonium: York Press.
- Flege, J., & Bohn, O. (2021). The revised speech learning model (slm-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3–83). Cambridge: Cambridge University Press. doi:10.1017/9781108886901.002.
- Gósy, M. (2004). *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.
- Hao, Y.-C. (2012). Second language acquisition of mandarin chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, 40, 269–279. doi:10.1016/j.wocn.2011.11.001.
- ’t Hart, J., Collier, R., & Cohen, A. (1990). *A perceptual study of intonation: An experimental phonetic approach to speech melody*. Cambridge: Cambridge University Press. doi:10.1017/CB09780511627743.
- Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70, 125–132.
- Juhász, K. (2023). Atonális és tonális nyelvek dallammeneteinek összehasonlítása. In *Alkalmazott Nyelvtudomány 2023* (pp. 21–46). (2nd ed.).
- Juhász, K. (2024). A mandarin beszédhangok produkciója kínaiul tanuló magyar anyanyelvűek ejtésében.

- Juhász, K. (előkészületben). Isolated mandarin chinese lexical tone production: a systemic approach analysing tone contrast within l2 learner groups.
- Klein, W. (1986). *Second language acquisition*. Cambridge: Cambridge University Press.
- Kuhl, P. (1980). Perceptual constancy for speech-sound categories in early infancy. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child phonology – perception* (pp. 41–66). New York: Academic Press.
- Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82, 1–26. doi:10.18637/jss.v082.i13.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Leather, J., & James, A. (1991). The acquisition of second language speech. *Studies in Second Language Acquisition*, 13, 305–341. doi:10.1017/S0272263100010019.
- Lee, Y.-S., Vakoch, D., & Wurm, L. (1996). Tone perception in cantonese and mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, 25, 527–542.
- Lenth, R. (2020). Emmeans: Estimated marginal means, aka least-squares means. URL: <https://CRAN.R-project.org/package=emmeans> letöltés:.
- Major, R. (2001). *The foreign accent: The ontogeny and phylogeny of second language phonology*. New York: Routledge. doi:10.4324/9781410604293.
- Markó, A. (2013). *Az irreguláris zöngé funkciói a magyar beszédben*. Budapest: ELTE Eötvös Kiadó.
- Markó, A. (2017). Hangtan. In A. Imrényi, N. Kugler, M. Ladányi, A. Markó, S. Tátrai, & G. Nagy (Eds.), *Nyelvtan* (pp. 75–203). Budapest: Osiris Kiadó.

- Mennen, I. (2014). Beyond segments: towards a l2 intonation learning theory (lilt. In E. Delais-Roussarie (Ed.), *Prosody and languages in contact: L2 acquisition, attrition, languages in multilingual situations*. New York: Springer. doi:10.1007/978-3-662-45168-7_9.
- Nolan, F. (2003). Intonational equivalence: an experimental evaluation of pitch scales. In *Proceedings of 15th ICPhS 2003* (pp. 771–774). Barcelona: Universitat Autònoma de Barcelona.
- Odlin, T. (1989). *Language transfer*. Cambridge: Cambridge University Press. doi:10.1017/CB09781139524537.
- Quené, H. (2014). hqmisc: Miscellaneous convenience functions and dataset. 0.1-1-es r csomagverzió.
- R Core Team (2024). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Letöltés*, . URL: <https://www.R-project.org/>. doi:10.32614/CRAN.package.hqmisc.
- van Rij, J., Wieling, M., Baayen, R., & van Rijn, H. (2020). itsadug: Interpreting time series and autocorrelated data using gamms. 2.4-es r csomag-verzió.
- Scarborough, R., & Zellou, G. (2013). Clarity in communication: "clear" speech authenticity and lexical neighborhood density effects in speech production and perception. *The Journal of the Acoustical Society of America*, *134*, 3793–3807. doi:10.1121/1.4824120.
- Shen, X. (1989). Interplay of the four citation tones and intonation in mandarin chinese. *Journal of Chinese Linguistics*, *17*, 61–74.
- Sóskuthy, M. (2021). Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *Journal of Phonetics*, *84*, 1–19. doi:10.1016/j.wocn.2020.101017.
- Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The lingu-

- istic proximity model. *International Journal of Bilingualism*, 21, 666–682. doi:10.1177/1367006916648859.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag. doi:10.1007/978-3-319-24277-4.
- Wood, S. (2017). *Generalized Additive Models: An Introduction with R*. New York: Chapman and Hall/CRC. doi:10.1201/9781315370279.
- Xu, Y. (1997). Contextual tonal variations in mandarin. *Journal of Phonetics*, 25, 61–83. doi:10.1006/jpho.1996.0034.
- Ye, Q. (2013). Pronunciation problems of hungarian university students learning chinese: Analysis, causes and possible solutions.
- Zhang, H., Wiener, S., & Holt, L. (2022). Adjustment of cue weighting in speech by speakers and listeners: Evidence from amplitude and duration modifications of mandarin chinese tone. *The Journal of the Acoustical Society of America*, 151, 992–1005. doi:10.1121/10.0009378.