

Grammatikaelmélet és kognitív pszichológia* Algebrai szabályok, statisztikai analógiák

FEHÉR KRISZTINA

1. A nyelvtudományban általánosan elfogadott tétel, hogy a grammatika szabályokból épül fel, és ezeknek megfelelően működik. A nyelvészeti munkákban másfajta nyelvtani szerkezet, illetve mechanizmus legfeljebb a kivételek kapcsán vetődik fel, de ekkor is kizárólag másodlagosan.

A szabályok per definitionem olyan logikai képletek, amelyek egyértelmű behelyettesítéseken alapulnak, ezért „minden vagy semmi” jellegűek, azaz kategorikusak. Egy nyelvi elemre egy szabály biztosan érvényes akkor, ha az az adott formula szerkezetével egybevágó, különben biztosan nem.

Mindazonáltal nincs grammatika, amiben ne lennének ebből a szempontból bosszantó kivételek. Olyan nyelvi adatok, amelyek megfelelnek a szabályok feltételeinek, de mégsem esnek a hatókörük alá, és olyanok, amelyek annak ellenére tartoznak egy adott szabályhoz, hogy az ehhez szükséges jegyeket maradtalanul hordoznák. Ezeket az eseteket a nyelvtanokban vagy egyszerűen a grammatikán kívülre helyezik, és listaszerűen megadható lexikális elemek sorának tekintik, vagy olyan rendhagyó szerkezeteknek tartják, amelyek háttérben a analógia áll. Utóbbin pedig azt a jelenséget értik, amikor az emberek grammatikai tudása „megtéved”, és egy-egy struktúrát mégsem a szabályoknak megfelelően, hanem egy fonetikájában-szemantikájában hasonló hangsor mintájára hoznak létre.

2. Ez a nyelvtankép hosszú ideje hagyományozódik a nyelvről való gondolkodásunkban. A grammatikai szabályok léte annyira magától értetődővé vált, hogy az elmúlt száz-százötven évben már szinte fel sem vetődött: volna más elvi lehetőség is, még hozzá egy olyan nyelvtan, amelynek egyetlen szervező elve az analógia.¹ Miközben a tapasztalat azt mutatja, hogy a grammatikákban a szabá-

* A kutatás a TÁMOP 4.2.4.A/2-11-1-2012-0001 azonosító számú Nemzeti Kiválóság Program — Hazai hallgatói, illetve kutatói személyi támogatást biztosító rendszer kidolgozása és működtetése konvergencia program című kiemelt projekt keretében zajlott. A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

¹ E tekintetben van ugyanakkor néhány kivételesnek mondható kezdeményezés is. A magyar szakirodalomban ilyenek például SZILÁGYI N., illetve a KÁLMÁN–REBRUS–TÖRKENCZY hármas

lyok mellett mindig számolni kell valamilyen rendhagyósággal (lexikai kivételekkel, analógiákkal), azt, hogy mennyire lenne összeegyeztethető az adatokkal egy teljes egészében analógiás nyelvtan, a nyelvtudományban korpuszokon módszeresen nem is igen tesztelték — noha a szabályalapú grammatikák kapcsán tisztázatlan kérdések, inkohereciát jelző *e l l e n t m o n d á s o k* mindig is adódtak.

Nem kézenfekvő például, hogy hogyan keletkezhetnek olyan lexikai kivételek, amelyeknek — mivel nem szabályosak — tulajdonképpen nincs is nyelvtani szerkezetük. Az analógiás rendhagyóság esetében pedig az nem világos, hogy ha az analógia az általánostól való egyedi eltérés csupán, az ilyen grammatikailag átstrukturálódott alakok miért terjedhetnek el, nem egyszer oly mértékben, hogy akár egész morfoszintaktikai csoportokat hozzanak létre. Ha viszont ezek az eltérések mégsem valamiféle sajátos asszociációkból, hanem abból adódnak, hogy az emberek kognitív nyelvtana szabályok helyett analógiákkal működik, akkor meg az válik kérdésessé, hogy evolúciósan honnan származhat az elsődlegesként kezelt szabálygrammatika, ha nem az emberektől. Mindezek mellett nem érthető még az sem, hogy egy ilyen rendszerben hogyan lehetséges egyáltalán változás — a szabályok merev osztályozást eredményeznek, így nem engedik meg a kategóriák közti mozgást, ezáltal mind az adatok átsorolódását, mind saját átértelmeződésüket kizárják. A nyelvtörténeti munkákban a változások kapcsán használt terminusok („grammatikalizáció”, „a tövéghangzók felső nyelvvállásúvá válása, eltűnése” stb.) háttérben sem valódi kauzális érvelések állnak: a magyarázatok nem az okokat, hanem a kiindulás és az eredmény viszonyát nevezik meg.²

3. A nyelvészek persze aligha hibáztathatók amiatt, hogy e problémák láttán sem gondolnak a grammatika modellezésének elméleti keretét érintő, átfogó reformokra. Az efféle *h o z z á l l á s* tudományfilozófiai nézőpontból voltaképpen természetes is. Ahogy WITTGENSTEIN írja: „A dolgoknak azokat az aspektusait, amelyek a legfontosabbak számunkra, egyszerűségük és mindennapiságuk rejti el előlünk. (Nem lehet észrevenni — mivel folyton szem előtt van.) Kutatásának tulajdonképpeni alapjai egyáltalán nem tűnnek fel az embernek. Hacsak ez nem tűnt fel egyszer neki. — És ez azt jelenti: az nem ötlük a szemünkbe, ami, ha egyszer észrevettük, a legszembeötlőbb és a legerősebb” (1952/1992: 83, a szövegbeli kiemelést töröltem).

azon újabb munkái, amelyek egy nem szabályalapú nyelvtan kidolgozását szorgalmazzák (SZILÁGYI N. 2004, 2013, KÁLMÁN 2007, 2012, KÁLMÁN–REBRUS–TÖRKENCZY 2010).

² A szabályok, illetve a szabály–analógia kettős problematikájáról nyelvelméleti-tudománytörténeti nézőpontból részletesen lásd FEHÉR 2013b, innen is különösen 68–72. Ugyanerről az újgrammatizmus és a nyelvi változások felől közelítve vö. még 2004, 2013c, a nyelvelsajátítás és a grammatika konnektionista modellje irányából 2008b, 2011b, ezen belül kifejezetten a morfoszintaxisra koncentrálva pedig 2013a.

Tanulmányomban arra törekszem, hogy a nyelvtudomány tradicionális grammatikaképéhez rendhagyó módon, immár empirikus tesztekre is építve, illetve némileg kívülről közelítsek. A szabályok és az analógia problematikáját ezúttal egy másik, a nyelv szerkezetével szintén foglalkozó, de a nyelvészeti hagyománytól kevésbé befolyásolt diszciplína, a kognitív pszichológia kísérleti méréseire támaszkodva fogom tárgyalni.

Ezzel mindenekelőtt az a szándékom, hogy úgy vessem fel egy analógia-grammatika elvi lehetőségét, hogy azt gyakorlati úton teszteljem is. Az érvek eredményeim szerint megerősítik majd egy ilyen nyelvtan létjogosultságát, de ahhoz mindenképpen hozzájárulnak, hogy az elmúlt másfél század szabályalapú grammatikáirása kapcsán felmerülő kérdéseket megválaszoljuk, és az ellentmondásokat feloldjuk.

4. A kognitív pszichológia számára a nyelvi képességek kutatása lényeges, ám nem az egyedüli feladat — ezt a területet lazább és másfajta szálak fűzik a lingvisztikai tradícióhoz, mint magát a nyelvészetet. A tudományág a pszichológiából vált önálló diszciplínává, és nem is oly régen. Miközben egyéb forrásokból is táplálkozott, a nyelvészetből szinte csak annak az ötvenes évek végétől induló formális grammatikai iskolának a megállapításait használta fel, amelyik a nyelv ügyét a mesterségesintelligencia-kutatáshoz is kapcsolódva, kifejezetten mint kognitív kérdést vetette fel.

Bár ez a nyelvtudományi hatás olyan értelemben markáns volt a kognitív pszichológia elméletére és módszertanára nézve, hogy ezt a tudományterületet a nyelv kapcsán jó darabig a szabályalapú koncepciók mellett kötelezte el, a formális nyelvészet koncepciója mégsem lett egyedüli irányadóvá. Az utóbbi időszakban az emberi elme evolúciós kontextusba helyezésével, a neurális hálózatok feltörésével, a gépi tanulás tanulmányozásának újabb eredményeivel, illetve a modellezési lehetőségek kitágulásával számos kritika és alternatívát jelentő elmélet is feltűnt a tudományterület szakmai diskurzusában. A szabályalapú megközelítés ugyan még itt is elterjedtebb, de már korántsem egyeduralkodó.

5. A kognitív pszichológiában a grammatika nyilvánvalóan mentális rendszerként értelmeződik. Ennek megfelelően kutatása a nyelvfeldolgozáson túl leginkább a nyelvvelsajátítás folyamatához kapcsolódik, és ez utóbbin belül is főként a köré a téma köré szerveződik, hogy egy-egy nyelv nyelvtanát vajon miféle tanulás útján fedezzük fel. Az álláspontok alapvetően két, egymástól élesen elkülönülő csoportba rendezhetők, a köztük lévő vita pedig jellegzetesen mint az algebrai és a statisztikai tanulás dilemmája fogalmazódik meg.

A nyelvi folyamatok kutatását a kognitív pszichológiában hosszú ideje az a CHOMSKY-féle felfogás (1957/2003, 1975, 1980) dominálta, hogy a nyelv jól formalizálható, logikai-matematikai jellegű műveletekkel pontosan leírható, a grammatika dekódolásához tehát ezeket az absztrakt szabályokat kell megtalál-

nunk. Amennyiben rájöttünk mondjuk arra, hogy az $y = x + 2$ egyenlet esetében az összefüggés x és y elemek között ezzel a képlettel írható le, akkor azt a jövőben bármelyik x -re tudjuk és fogjuk alkalmazni — innen az „algebrai tanulás” elnevezés.

Egy az ezen a területen híres nyelvi példával élve: általános szintaktikai szabályként tartják számon, hogy az angolban bármely plurális főnévi frázis és bármely vele számban egyeztetett igei frázis összefűzésével alkothatunk mondatot. Ha ismerjük ezt, továbbá felfedezzük, hogy a *blickets* (ami egyébként egy teljesen nem létező szó) egy plurális főnév kell legyen, akkor azt is tudjuk, hogy a *the three blickets* és a *reminded Sam of Tibetan art* kifejezések egy-egy egymással számban egyeztetett főnévi és igei frázisnak tekinthetők. Ebből pedig egyúttal arra is következtetünk, hogy a *the three blickets reminded Sam of Tibetan art* egy jól formált angol mondat (MARCUS és mtsai 1999: 77–78).

Ennek a gondolatmenetnek az ellentételezéseként tűnt fel a statisztikai koncepció, amely viszont leginkább úgy jellemezhető, mint egy tudásbővülési folyamat arról, hogy egy korpuszban az egyes elemek milyen eloszlásban találhatók meg a többihez viszonyítva. Statisztikai tanulás esetében absztrakt összefüggések helyett ún. átmenet-valószínűségeket keresünk a struktúra megértéséhez — olyanokat, mint hogy mondjuk a **ba* megjelenésének esélye **pi* előtt 0,042, **ko* után 0,005, **ta*-t másodikként követve pedig 0,013.³ Ez utóbbi mérőszámok azt mutatják meg, hogy egy-egy adott elem mekkora eséllyel jósol meg egy valamelyik másik, szomszédos vagy közvetve következő elemet. A $P(B|A)$ átmenet-valószínűség tehát annak az esélye, hogy B jön az A után. Fontos, hogy ez nem ugyanaz, mint annak a valószínűsége, hogy egy korpuszban AB jelenik meg: az AB kombináció előfordulási gyakorisága például lehet igen kicsi, miközben az A-ból B-be való átmenet-valószínűség nagyon nagy (ez akkor áll fenn, ha A maga is ritka elem a korpuszban, de amikor feltűnik, rendre vagy majdnem mindig B követi) (vö. LUKÁCS 2014: 524).

Egy a kognitív pszichológiában megintcsak közismert nyelvtani példát használva: a *The boy loves apples.* és a *The boy loves oranges.* mondatokból álló, a mondatok előfordulása tekintetében kiegyensúlyozott korpuszban annak a valószínűsége, hogy a *the* és a *boy*, valamint a *boy* és a *loves* szavak közvetlenül egymás után és ilyen sorrendben jelennek meg, 1,0, de a *loves* és *apples*, illetve a *loves* és *oranges* párok esetében ugyanez már csak 0,5, a **the loves*, **the apples*, **the oranges*, **boy apples*, **boy oranges* stb. kapcsolatok előfordulásának esélye pedig 0,0. Ha a korpusz láttán-hallatán folyton ezeket a statisztikai mutatókat tapasztaljuk, egy idő után tudni fogjuk, hogy ebben a nyelvben csak ez a kétféle grammatikus mondat lehetséges (vö. MARCUS és mtsai 1999: 77).⁴

³ A * jellel itt egy kitalált nyelv elemeire, alább pedig egy feltételezett korpusz nulla előfordulású adataira utalok.

⁴ Mindehhez lásd még FISER 2005: 26–29.

6. Könnyen meglehet, hogy az algebrai és a statisztikai tanulás eltérése egyben a nyelvtani rendszerek különbsége is. Ez abban az esetben van így, ha vizsgálódásaink során azt látjuk, hogy a gyerekek nyelvelsajátítása, illetve a kognitív nyelvi folyamatok általában véve azért vagy azért is zajlanak egy adott mechanizmus szerint, mert az emberi nyelv grammatikája szerkezetéből adódóan más-képp nem is volna megfejthető és működtethető. Erről az oldalról közelítve tehát az algebrai–statisztikai grammatikatanulás dilemmája a kognitív pszichológiában megegyezhet, de legalábbis szorosan összefügg a szabály–analógia problematikával a nyelvészetben.

Ebből a szemszögből a szabály voltaképpen egy olyan absztrakt képlet, amibe a neki maradéktalanul megfelelő nyelvi elemek helyettesítendők be. Ha pedig az analógiát immár ez utóbbi viszonylatában nézzük, akkor a szokásos nyelvtudományi jellemzése helyett sokkal inkább azt a tulajdonságát érdemes kiemelni, hogy a szabályokkal ellentétben nem kategorikus, hanem skálaszerű értékekkel jellemezhető. Lényegében egy olyan valószínűségi mintázat, aminek az egyes nyelvi elemek logikai egyezések helyett hasonlóságuk foka szerint, vagyis különböző mértékben, statisztikai alapon feleltethetők meg.

7. Mindezeket figyelembe véve az alábbiakban a kognitív pszichológiában manapság is igen éles „algebrai vs. statisztikai tanulás” vitát ismertetem, mégpedig a legjelentősebbnek tartott kísérletek rétamaszkodva. Ez ugyanakkor nem pusztán egy másik diszciplína vonatkozó szakirodalmának kissé iskolás összefoglaló bemutatását jelenti nyelvészek számára — áttekintésem ennél legalább két szempontból többnek szánom.

Amellett, hogy a tanulságokat a nyelvtudomány szabály–analógia párjára vonatkoztatom, egyes teszteket át is értelmezek. A célom éppen az, hogy a kognitív pszichológia empirikus vizsgálatai segítségével ne pusztán felvessem egy statisztikai grammatika elvi lehetőségét, de példákkal is szemléltessem annak gyakorlati működőképességét. Ezért azoknak a pszicholingvisztikai kísérleteknek, amelyeknek az adatait az algebrai tanulás melletti érvként szokták felhozni, igyekszem olyan új elemzést adni, amely rávilágíthat az adatok eredeti értelmezésének módszertani hibáira és arra, hogy kiküszöbölésükkel a mérési eredmények a szabálygrammatikának nem feleltethetők meg, ám egy statisztikai nyelvtannak igen.

8. A kognitív pszichológiában az „algebrai vs. statisztikai tanulás” vitát tulajdonképpen SAFFRAN, ASLIN és NEWPORT egyik kísérlete (1996) robbantotta ki. Ez elsőként próbálta meg bemutatni, hogy a nyelvi struktúrák elsajátításának mindaddig szinte egyeduralgó, algebrai koncepciója mellett létezik egy másik, statisztikai forgatókönyve is.

Tesztjükhez egy olyan műnyelvet konstruáltak, amelynek grammatikája egyetlen dologgal volt jellemezhető, mégpedig az elemek kapcsolódási valószínűségével (A változat: *bidakupadotigolabubidaku...*, illetve B változat: *dapikutilado-*

burobidapiku...). Ennek négy szava tizenkét szótaghármából állt elő úgy, hogy a szótagok kombinálódási esélye előre meghatározott értékeket követett (A eset: *golabu, bidaku, padoti, tupiro*, valamint B eset: *dapiku, tilado, burobi, pagotu*).

A kitalált szavakból egy kétperces szöveget állítottak össze, amelyben minden szóból negyvenötöt szerepeltettek, és ugyanaz a szó nem fordulhatott elő kétszer egymás után. Ezt az arányossági ellensúlyozást leszámítva azonban a szavak megjelenése teljesen random módon történt. Az egy egységet képező elemsorok (sztringek) belső szerkezete ugyanakkor már korántsem volt véletlenszerű: az egy szóba tartozó elemek mindig egy bizonyos sorrendben követték, ezáltal tökéletesen megjósolták egymást. A szavakon belüli szomszédos szótagok közti átmeneti valószínűség 1,0 volt, a szó végi és eleji elemek között csak 0,3, míg a szó közepi szótagok esetében, amelyek sose lehettek szomszédosak, 0,0. Tehát például az A verzióban *tu* után mindig *pi*, majd *ro* következett (1,0), a B változatban például *tu* után előfordulhatott *da, ti* és *bu* is (0,3), az A verzióban pedig *ro* után nem állhatott *pi, la* vagy *do* (0,0).

A kísérletvezetők ügyeltek arra, hogy ezeken a statisztikai információkon túl ne legyen semmilyen más fogódzó, amiből a szóhatárok helyére lehetne következtetni. A szöveget beszédszintetizátorral állították elő, így egy monoton női hangot lehetett hallani, ami percenként átlagosan kétszázhetven koartikulált szótagot produkált, szünetek nélkül, kiegyenlített szupraszegmentális jegyekkel.

9. SAFFRANÉK ezt a mesterséges nyelvet az ún. familiarizációs (megismertelési), habituációs (szoktatási) vagy tréning- (tanító)fázisban nyolc hónapos csecsemőkkel hallgattatták meg: két csoportra osztották őket, és az egyikkel az A változatot, a másikkal a B verziót ismertették meg. Rögtön utána egy másik hanganyagot is mutattak nekik, hogy értékelhessék, a babáknak sikerült-e rájönniük, melyek az előzőleg bemutatott nyelv szavai.

Az újabb szövegben az A változathoz vett *tupiro* és *golabu*, valamint a B verzióból származó *dapiku* és *tilado* ingereket (stimulusokat) ismételték, mégpedig mindegyiket háromszor, egyforma szüneteket tartva közöttük. Ebben a szakaszban tehát minden egyes csecsemőnek összesen tizenkét tesztelendő példánya (tokenje) volt. Ennek négy fajtájából (típusából) kettővel már találkoztak a korábban lejátszott hanganyagban, ezek az általuk hallott felvétel szavai voltak, míg kettő ismeretlen volt számukra, ezek az előző szövegben nem voltak szavak.

A vizsgálat során SAFFRAN és munkatársai a csecsemőknek arra a tulajdonságára alapoztak, hogy a babák a megszokott és az új ingereket eltérő mértékű érdeklődéssel fogadják. Ennek meghatározásához az ún. „fejfordítási preferencia paradigmát”⁵ alkalmazták, és azt találták, hogy a nyolc hónaposok mindkét cso-

⁵ A „fejfordítási preferencia paradigma” („HPP”, azaz „Headturn Preference Paradigm”) egy olyan kísérleti protokoll, amelynek során a sztringeket oldalirányból, fényel kísérve játsszák le, és azt figyelik, hogy milyen ingerek hallatán és mennyi időre fordítja a lámpás hangszóró irányába a baba.

portban szignifikánsan jobban odafigyeltek azokra a kombinációkra, amik nekik újak voltak. Mindez pedig annak a jele, hogy felfedezték a tanító szövegben a szóhatárokat, vagyis elszakították a szóhangalakokat, ezáltal pedig képesek voltak felismerni a kétféle szótaglánc közti strukturális különbséget.⁶ Olyannyira, hogy amikor a mérést nehezített feltételekkel megismételték, akkor is hasonló eredményt kaptak.

Csecsemők további két csoportjának is lejátszottak egy-egy újabb, az előbbihez hasonló tanító szöveget. Az A nyelv szavai most a *pabiku*, *tibudo*, *golatu* és *daropi* voltak, a B-é pedig a *tudaro*, *pigola*, *bikuti* és *bodupa*. Am ezúttal a babákkal a tesztfázisban olyan stimulusokat hallgattattak, mint a *pabiku*, *tibudo*, *tudaro* és *pigola*. Ezek közül az első kettő az A változatban, a második kettő a B verzióban volt szó. Azok a szótaghármak pedig, amelyek nem voltak szavai az adott nyelvnek, részzavak voltak, vagyis olyan szótagsorok, amelyeket egy szó utolsó szótagjából és egy másik szó első két szótagjából állítottak össze (az A verzióban részzónak minősülő *pigola* például az A nyelvbeli *daropi* és a *golatu* szavakból állt össze).

Miután ez esetben a nemszavak is olyanok voltak, amelyeket hallhattak a csecsemők a tanító szövegben, a feladat korántsem volt könnyű, a babák mégis jól teljesítettek. Az előző teszttel szemben, ahol a nemszavak első és második, illetve második és harmadik szótagjai között 0,0 volt az átmenet-valószínűség a tanító szöveghez viszonyítva, most olyan szótaghármak között kellett tudniuk különbséget tenni, amelyeknek két szomszédos szó részletéből előálló nemszavaiban a szótagok között nem lehetett 0,0 az átmenet-valószínűség. Most az átmenetesély értéke a szavakban a szomszédos szótagok között 1,0 volt, míg a részzavakban, az első két szótag között 0,3, a második és harmadik szótagok közt pedig — hasonlóan a szavakhoz — 1,0.⁷

10. Noha maga a mérés igen meggyőző, a kísérletvezetőknek azt a kijelentését, amelyben a nyelvi szerkezetek elsajátítása kapcsán a statisztikai mechanizmusok jelentőségét hangsúlyozták, többen is kétségbe vonták. SAFFRANÉK cikke nagy ellenállást váltott ki, számos új teszt született, amely hasonló módszerekkel éppen az *algebrai* mellett kívánt érvelni.

⁶ A teszt rövid bemutatása magyarul FISER 2005: 28–29, PLÉH–LUKÁCS–KAS 2008: 843, illetve MACWHINNEY 2003: 507, TAUZIN 2010: 390; vö. továbbá FISER 2005: 26. Újabb megerősítő empirikus adatokhoz vö. SAFFRAN és mtsai 1997, ASLIN–SAFFRAN–NEWPORT 1998, PERRUCHET–PEEREMAN 2004, WHITE és mtsai 2008, PELUCCHI–HAY–SAFFRAN 2009a, 2009b. Ezen belül is a csecsemők viselkedését nem humán fajok egyedeivel összevetve lásd SAFFRAN és mtsai 2008, ennek más kísérletekkel társított elemző ismertetését TAUZIN 2010: 392–393. A szóhatárok detektálásának neurolingvisztikai korrelátumaihoz pedig vö. SANDERS–NEWPORT–NEVILLE 2002, SANDERS–NEVILLE 2003, CUNILLERA és mtsai 2006, ABLA–KATAHIRA–OKANOYA 2008, BUIATTI–PEÑA–DEHAENE–LAMBERTZ 2009, további vonatkozó fMRI-vizsgálatokhoz lásd még TURK–BROWNE és mtsai 2008.

⁷ Magyar ismertetését lásd LUKÁCS 2014: 525.

E tekintetben az első nagyobb hatású munka a MARCUSÉKÉ (1999), akik azt vizsgálták meg, hogy ha teszteléskor egy adott struktúrát csupa olyan elemmel prezentálnak, amiket a szoktatás során egyáltalán nem használtak, a babák ezek közül is kiszűrnek-e a korábban hallottakkal szerkezetileg egyező szekvenciákat. Azt gondolták, hogy ha a csecsemők így is képesek differenciálni, teljesítményük statisztikai mechanizmussal nem magyarázható, a tanulás ez esetben algebrai szabályok kinyerése révén kell, hogy történjen.

Beszéd szintetizátorral kétféle, szupraszegmentális jegyei tekintetében kiegyenlített tanító szöveget készítettek, amit hét hónaposoknak két percen át játszottak le: a babákat — csakúgy, mint SAFFRAN és kollégái (1997) — két csoportra osztották, és az egyik csapatot az egyik hanganyaggal, a másikat a másikkal ismertették meg. Mindkettőben tizenhat háromszavas mondat volt hallható, mégpedig úgy, hogy a mondatok között egy, ezek szavai között pedig negyed másodperces szüneteket tartottak. Mindegyik mondat háromszor ismétlődött, egymáshoz képest random módon. A bemutatott minták nyelvtana egy általános képletet követett: a mondatok egyik esetben az ABA formulának (*ga ti ga, li na li, ni gi ni* stb.), a másik változatban az ABB struktúrának (*ga ti ti, li na na, ni gi gi* stb.) feleltek meg. Ezek a mondatok mindkét verzióban négyféle ismétlődő és egyszeri elemet tartalmaztak: ami az ABA-ban repetitív volt, az az ABB-ben nem (*ta, ga, ni, li*), ami pedig az ABA-ban nem ismétlődött, az az ABB-ben igen (*ti, gi, na, la*).

11. Ezután jött MARCUSÉK mérésének tesztfázisa, amelyben a babák négy mondatot hallottak véletlenszerű sorrendben, mindegyiket háromszor ismételve. Ezek közül kettő (*wo fe wo, de ko de*) az ABA, kettő (*wo fe fe, de ko ko*) pedig az ABB nyelvtanhoz illeszkedett. Mindez a csecsemők számára a gyakorlatban azzal járt, hogy ebben a hanganyagban egyformán tűntek fel a tréning szöveg szempontjából grammatikus és agrammatikus mondatok: az ABA nyelvtannal trenírozott babáknak az ABB szerkezetűek jelenthettek újdonságot, és fordítva.

MARCUSÉK kísérletükben SAFFRAN munkacsoportjához (1996) hasonlóan a szoktatásból adódó preferencia mérésére alapoztak, amihez ugyancsak a fejfordításos paradigmát használták. Az adatok értékelésekor a csecsemők figyelmi reakciója azt mutatta, hogy a babák az általános struktúrák között is képesek differenciálni, mivel — függetlenül attól, hogy a tanító szövegnek melyik változatát hallották — a tréning szakasz nyelvének grammatikájához illeszkedő formuláknál láthatóan nagyobb érdeklődést mutattak az agrammatikus szerkezetek iránt. Ebből pedig a kutatók arra a következtetésre jutottak, hogy a csecsemők ez esetben a l g e b r a i ú t o n tanultak, hiszen a megismertetés során nem szerezhettek a tesztfázis mondatairól konkrét tapasztalatokat.⁸

⁸ A kísérlet magyar nyelvű összefoglalását lásd még FISER 2005: 29, 31–32, LUKÁCS 2005: 39–40, TAUZIN 2010: 390–391. E feltételezett kognitív eljárás rekurzívan hierarchikus jellegéhez FITCH-HAUSER 2004, az eredményeket a nem humán fajok esetében tesztelő vizsgálatok elemző bemutatásához pedig TAUZIN 2010: 397–401.

12. Jóllehet logikusnak tűnik, olyan konklúzió, miszerint a hét hónapos babák nyelvtan-elsajátítása algebrai tanulással történik, MARCUSék vizsgálatából — annak sajátos módszer t a n i k ö r ü l m é n y e i folytán — mégsem vonható le. A teszt interpretációja során ugyanis legalább két dolgot nem szabad figyelmen kívül hagyni.

Nem lehet eltekinteni attól, hogy a tréning- és a tesztfázis mondatait alkotó elemek eloszlási mutatói között több kontrollálatlan megfelelés is volt. Ezek mennyiségi és minőségi szempontból egyaránt torzíthatták a kísérleti adatokat. Azon túl, hogy plusz információkként segíthették a csecsemőket a struktúrák azonosításában, algebrai tanulásnak tüntethették fel azt is, ami esetleg valójában nem vagy nem teljesen volt az. Ezen tényezők közül kettőt egyébként maguk a vizsgálatot végző szakemberek is megemlítenek cikkükben.

Az egyik ilyen a szoktató és a tesztelő stimulusok hangalaki jellemzőinek korrelációjából adódott. A tesztszövegben ugyan nem szerepelt a tréningfázis egyetlen példája sem, a teszt grammatikus és agrammatikus mondatai mégsem pusztán olyan értelemben mutattak összefüggést a tanító anyaggal, hogy követték-e az ABA, illetve az ABB formulákat, hanem zöngességi szerkezetük tekintetében is. Míg a két grammatikus tesztmondat ilyen szempontból a szoktatás során használt típusok közül hárommal is megegyezett, a két agrammatikusé egygyel sem. Ennek következtében előfordulhatott, hogy a babák nem az általános formulát, hanem a mondatokban szereplő szavak zöngességi mintázatát memorizálták, és később ennek alapján tettek különbséget a grammatikus és agrammatikus mondatok között.

Az ABA nyelv szempontjából a teszteléskor a nyelvtant a *wo fe wo* és a *de ko de* követte, a *wo fe fe* és a *de ko ko* nem, az ABB oldaláról nézve pedig éppen fordítva. Miközben az ABA nyelv megismertetésekor a csecsemők több olyan típust is hallottak (*ga ti ga, li ti li, ni ti ni*), amelynek kezdő konzonánsai éppúgy a [+zöng] [-zöng] [+zöng] sorrendben jelentek meg, mint a tesztszöveg két grammatikus mondata (*wo fe wo, de ko de*), a habituáció során nem találkoztak egyetlen olyan példával sem, amely a teszteléskor használt anyag agrammatikus esetei (*wo fe fe, de ko ko*) [+zöng] [-zöng] [-zöng] mássalhangzó-szerkezetének felelt volna meg. Mindez elmondható az ABB nyelvről is: azok a babák, akikkel ezt a nyelvtant ismertették meg, olyan tokeneket nem hallhattak a tréningszakaszban, amelyek a *wo fe wo* és *de ko de* agrammatikus mondatok [+zöng] [-zöng] [+zöng] konzonánsstruktúráját mutatták volna, olyan szekvenciákat (*ga ti ti, li ti ti, ni ti ti*) viszont lejátszottak nekik, amelyek a grammatikus *wo fe fe* és *de ko ko* [+zöng] [-zöng] [-zöng] mássalhangzó-eloszlását követték.

Mindezek miatt MARCUS és kollégái elvégeztek egy módosított kísérletet is, amelyben a habituáció idején csak zöngés konzonánsokat tartalmazó típusokat (az ABA nyelvtan esetében például a *le di le, le je le, le we le* szókombinációkat) használtak, míg a tesztelésnél zöngétleneket is, mégpedig zöngésekkel variálva (*ba po ba, ko ga ko, ba po po, ko ga ga*).

Miután az eredmény megegyezett az előzővel, ebből a szempontból az ellenőrző vizsgálat sikeresnek tekinthető. Ez ugyanakkor már nem mondható el a másikkal, a kísérletvezetők által is felismert zaj kiiktatásáról, ami ráadásul éppen az általános formulákkal volt kapcsolatos. Az a tény, hogy az ABA-ban nincs, az ABB-ben pedig van közvetlen elemismétlődés (reduplikáció), lehetőséget adott a csecsemőknek arra, hogy a tréningfázis mondatainak struktúráját egy egyszerű kvantitatív megfigyeléssel fedezhessék fel. Bár MARCUS és kollégái elvégeztek egy olyan kiegészítő tesztet, amiben az AAB és az ABB reduplikatív formulákról mutatták meg, hogy a csecsemők ezek között is ugyanúgy differenciálnak, a problémát ez ténylegesen nem orvosolta. Az ismétlődés helyének („dupla után szimpla” vagy „szimpla után dupla”) eltérése alapján, azaz statisztikai információkra hagyatkozva a két struktúra megintcsak könnyen elkülöníthető volt.

Mindezeket túl nem lehet elfelejtenni arról sem, hogy a kísérletvezetők a familiarizációs szakaszban SAFFRANÉKTÓL (1996) eltérően jártak el: a mondatokat nem egyetlen fluens beszédláncban, hanem tagolt szóhármakban jelenítették meg. Ez a szegmentáció nemcsak megkönnyítette a babák dolgát, ezáltal pedig növelhette teljesítményüket, de annak is fokozta az esélyét, hogy a csecsemők a stimulusok nyelvtanát absztrakt változók kiemelése nélkül, pusztán a szünetek és a mondatok szavainak relatív előfordulási sorrendjére hagyatkozva fedezzék fel. MARCUS és kollégái tehát a tanulás algebrai jellegét úgy kívánták demonstrálni, hogy közben vizsgálatuk körülményei a statisztikai megismerési mechanizmust nem zárták ki (FISER 2005: 27, 31–32).⁹

13. A vitába bekapcsolódó PEÑA és csoportja (2002) érdekes módon előfeltételezte, hogy *m i n d k é t t a n u l á s i f o r m a* létezik. Így a kérdés szerintük valójában nem is az, hogy a kettő közül melyik stratégiával kell számolnunk, hanem az, hogy mikor és milyen mértékben érvényesül az egyik és a másik. A kutatók az eddig ismertetett kísérletek kapcsán ennek megfelelően leginkább azt hiányolták, hogy a statisztikai és az algebrai tanulásról olyan mérések alapján folyik vita, amelyek közvetlenül nem is vethetők egybe. Meglátásuk szerint a SAFFRAN- (1996) és a MARCUS-féle (1999) tesztek eléggé különbözőek, ezért összehasonlításukkal valójában nem lehet megítélni a kétféle tanulási mechanizmus egymáshoz viszonyított jelentőségét a kognitív folyamatokban.

A kutatók annak érdekében, hogy a méréseket ilyen szempontból egyértelműsítsék, egy teszt sorozatot terveztek. Ebben — nem melleleg éppen NEWPORT és ASLIN javaslatára (vö. PEÑA és mtsai 2002: 604) — más típusú mesterséges szókkal, az AXC struktúrával dolgoztak, miközben az eredmények függvényében folyton változtatták a vizsgálati körülményeket, de mindig csak egyet-egyet.

Az AXC nyelvnek a grammatikája úgy írható le, hogy az A megjósolja a C-t, ám nem jelzi előre az X-et. PEÑÁÉK azért választották éppen ezt az általános for-

⁹ Mindehhez vö. még LUKÁCS 2005: 40, TAUZIN 2010: 391, LUKÁCS–GÁBOR és mtsai 2014: 514.

mulát, mert úgy gondolták, ez a szerkezet alkalmas arra, hogy a két kognitív folyamatot elkülönítetten, de összevethető módon teszteljék. Érvelésük szerint ha az X helyén a habituáció során is használt elemeket jelenítenek meg, előfordulási valószínűségekkal definiálható, vagyis statisztikai információk révén elsajátítható formulákat kapnak, ha azonban második szótagokként a szoktatáskor nem tapasztalt hangkapcsolatok szerepelnek, a szekvenciális összefüggések csak elvont szabályokra támaszkodva tanulhatók meg.

PEÑÁék az ötrészes kísérletsorozatukhoz egy kilencszavas nyelvet hoztak létre három A–C párból és három X-ből. Ezeket az AXC struktúrákat egy beszéd-mintában úgy rendezték el, hogy a szövegbeli átmeneti valószínűségek a következők legyenek. Az A_iX , valamint az XC_i szóbelseji kombinációk előfordulási esélye 0,33, a szóhatárokon megjelenő C_iA_j kapcsolatoké 0,5, míg az egy-egy X-et körülvevő, szavakon belüli A_i-C_i pároké rendre 1,0.¹⁰ Mindez a gyakorlatban azt jelentette, hogy az A_1XC_1 szócsalád tagjai a *puliki, puraki, pufoki* voltak, az A_2XC_2 -é a *beliga, beraga, befoga*, az A_3XC_3 -é pedig a *talidu, taradu, tafodu*.

Ezt a nyelvet ezúttal nem csecsemőknek, hanem francia anyanyelvű felnőtteknek játszották le. A tanító szöveget az eddigi gyakorlathoz hasonlóan egy női hang beszédéből szintetizátorral hozták létre: a mintában az *r*-ek — igazodva a vizsgálatban részt vevők nyelvének fonetikájához — uvulárisak voltak, viszont a szöveg szemantikailag üres volt, a dallam-, hangsúly- és ritmusjegyek tekintetében pedig homogenizálták. A hanganyagot először szünetekkel sem tagolták, ily módon tehát a kutatók szerint nem tartalmazott semmilyen plusz információt, amire az AXC szavak azonosításában lehetett volna támaszkodni.

14. A tízperces habituáció után az első tesztfázis következett. Ebben a szakaszban a kísérleti személyekkel sztringpárokat hallgattattak. Egy-egy, a tanító szövegben is szereplő A_iXC_i szó mellett egy-egy olyan, az AXC nyelv szempontjából nemszót mutattak be, amely két szomszédos AXC szó részletéből állt össze (C_kA_iX : *kitara, kitafo, gapufo, dubera*; valamint XC_iA_j : *likita, lidube, radube, ragapu*), de természetesen nem esett egybe egyetlen létező francia szó hangalakjával sem. Miután a kutatók arra kérték a vizsgálatban részt vevőket, ítélik meg, hogy a kettősökből melyik tűnik inkább olyannak, mint a korábban lejátszott nyelv szavai, az alanyok szignifikánsan többször választották az AXC formulának megfelelő szavakat.¹¹

¹⁰ Az általános formulák A és C elemeinek alsó indexei a matematikából és a nyelvészetből is ismert hagyománynak megfelelően itt is a koreferenciaviszonyokat jelölik, vagyis azt, hogy egy adott A egy bizonyos C-vel ugyanazon vagy különböző szavak részeit alkotják-e.

¹¹ Szükséges megemlíteni, hogy PEÑÁ és munkatársai a stimulusok összeállításakor még két kontrollpróbát is beiktattak a vizsgálatok sorába, mert úgy gondolták, előfordulhatott, hogy a tanító és a tesztelésre használt hanganyagok bizonyos jegyei torzították az eredményeket. Egyrészt meggyőződtek arról, hogy a teszt nemszavainak a familiarizációs szövegbeli ritkább megjelenése nem befolyásolta-e a válaszadást. Másrészt leellenőrizték azt is, hogy nem volt-e hatással a reakciókra az az esetleges ritmus sem, amit az alanyok abból adódóan érezkelhettek, hogy az AXC szavak és

PEÑÁék az első sorozatot követően egy újabb, második mérést terveztek, hogy eldönthessék, a kísérleti személyek az első vizsgálatban produkált tudásukat miféle tanulással szerezték meg. Ehhez francia felnőttek egy másik csoportjának ugyanazt a beszédmintát játszották le, mint az előző csapatnak, de a tesztlésre használt ingeranyagot egy tekintetben módosították.

A teszt párok egyik tagjaként megint a $C_k A_i X$ és $X C_i A_j$ szerkezetű nemszavak szerepeltek, miközben másik tagként csak $A_i X * C_i$ felépítésű szók jelentek meg. Ezek olyan szavak voltak, amelyek követték az AXC nyelvtant, ám az alanyok korábban nem hallhatták őket, mivel az X-eik helyére új, a másik két szócsalád A vagy C típusú szótagjait tették.¹² Tudván, hogy a habituációs szöveg egyik szócsaládjá a *puliki, puraki, pufoki* volt, egy másikban pedig a *beliga* tűnt fel, PEÑÁék például arra kérdezték rá, hogy a résztvevők a *pubeki* vagy a *likibe* szótaghármast részesítik-e előnyben. Az előbbi szó ugyan, de az alanyok eddig nem találkoztak vele, az utóbbi az AXC nyelv szempontjából nemszó, viszont a kísérleti személyek a familiarizációs szakaszból ismerhetik.

Mivel a tesztnek ebben a változatában a résztvevők teljesítménye a felére esett vissza, PEÑA és munkatársai a következőkben huszonöt milliszekundumos szüneteket iktattak be a habituációs szótaghármassok közé. Ezzel a harmadik méréssel azt kívánták megvizsgálni, vajon a hanganyag részleteinek akár a perceptiós küszöbérték alatti (azaz nem is tudatosuló) tagoltsága befolyásolja-e az általános formulák felismerését.¹³ Ezt a tanító szöveget használva az előbbi tesztet tehát egy újabb csoporttal is elvégezték.

Kiderült, hogy ilyen familiarizációs feltételek mellett a felnőttek képesek arra, hogy még a soha nem hallott X^* elemek esetében is szignifikánsan többször válasszák az AXC szerkezetnek megfelelő hangsorokat ismerős kombinációként. A kutatók mindebből azt a konklúziót vonták le, hogy valóban *k é t f é l e t a n u l á s i m e c h a n i z m u s* megosztott jelenlétével kell számolni. Míg a statisztikai módszer fluens beszédben, az algebrai stratégia tagolt szöveg esetén érvényesül.

PEÑA és kollégái úgy gondolták, ez a következtetésük már csak azért is megalapozott, mert azt még két további kísérletük eredménye is megerősíti. Ezek

nemszavak szótagjai perceptuális középpontjainak egyenetlen lett az eloszlása. Mivel mindkét mérés a korábbival egyező eredményekkel zárult, a kutatók úgy vélték, a stimulusokat kellő módszertani gondossággal válogatták össze.

¹² Az általános formula képletében a * jel ennek megfelelően itt és a továbbiakban az előzőleg az adott pozícióban nem tapasztalt szótagot jelöli.

¹³ Azt, hogy ennek az újabb beszédmintának a szünetei olyanok, amelyek a perceptió során valóban nem tudatosulnak, egy kontrollméréssel ellenőrizték. Egy csoportnak ezt a familiarizációs hanganyagot mind folyamatosan, mind pedig a szavak között szüneteket tartva lejátszották, és megkérdezték tőlük, tartalmaz-e valamelyik megszakításokat. Miután az alanyok nemmel válaszoltak, elárulták nekik, hogy az egyik szövegben voltak szünetek. Amikor pedig ezt követően arra kérték őket, ennek tükrében mondják meg, melyikben, a válaszok véletlenszerűnek bizonyultak.

megtervezésekor a kutatók abból indultak ki, hogy amennyiben a tanulás algebrai, illetve statisztikai jellege tényleg az ingeranyag tagoltságán múlik, a habituációs időtartamát megváltoztatva a következőre lehet számítani.

Egyrészt az első és második kísérletben használt beszédminta jelentős növelése sem fog absztrakt formulák felismeréséhez vezetni, ha az ömlesztetten lejátszott hanganyag esetében algebrai tanulás helyett statisztikaival számolhatunk: az alanyok a szoktatás során nem tapasztalt X^* -ek hallatán továbbra sem fognak jól teljesíteni a tesztben. Másfelől viszont az általános struktúrák kiemelése még a harmadik mérés familiarizációs szakaszának drámai csökkentése mellett is sikeres kell, hogy legyen, ha a szünetekkel elválasztott szótaghármak hallgatásakor nem statisztikai, hanem algebrai tanulás történik: a kísérleti személyek az oktató anyagban nem szereplő X^* -ekkel találkozva ekkor is jól fognak teljesíteni a tesztben.

Mindezeknek megfelelően egy negyedik kísérletben a tagolatlan szótaghármak beszédmintájának időtartamát a háromszorosára emelték, egy ötödik vizsgálatban pedig a tagolt sztringek tréningiszövegének bemutatási idejét ötödére csökkentették, majd ezek után mindkét esetben a harmadik kísérlet során használt tesztingerekkel ellenőrizték a tanulás eredményét. A kísérleti személyek a várt módon teljesítettek: ömlesztve bemutatott szótaghármak harmincperces szoktatási fázisa után sem ismerték fel az új X^* -eket tartalmazó grammatikus szerkezeteket, miközben csupán egy kétperces, a szavakat tagoltan találó habituációs szakaszra alapozva képesek voltak arra, hogy soha nem hallott X^* -ek esetében is megtalálják azokat a sztringeket, amelyek az AXC nyelvtant követték.¹⁴

15. A PEÑÁék által elvégzett vizsgálatsorozat ugyan módszerben már lényegesen jobban kontrollált, mint a MARCUSÉké (1999) volt, mégsem mondható kifogástalannak.

A harmadik és ötödik kísérlettel kapcsolatban például mindjárt felvethető problémaként, hogy az alanyok a szüneteket mint önálló elemeket kezelhették a hangsorfolyamban. Így ezekre hagyatkozva valójában statisztikai összefüggéseket fedezhettek fel, majd pedig ennek segítségével oldhatták meg a tesztfeladatot. Erre az eshetőségre maguk a kutatócsoport tagjai is gondoltak, ezért egy-egy kontrollkísérlettel igyekeztek álláspontjukat megerősíteni.

A harmadik mérés egy újabb verziójában PEÑA és munkatársai a C_kA_iX és XC_iA_j szerkezetű nemszók szótaghármakainak C és A elemei közé is szüneteket helyeztek. Ezzel azt kívánták megfigyelni, hogy változtat-e az eredményeken, ha a tagolás nem mutat egyértelmű megfelelést a szóhatárokkal. Az ötödik kísérlet ellenőrzése során pedig azt igyekeztek kiküszöbölni, hogy a kísérleti személyek arra a statisztikai korrelációra alapozzanak, hogy az A-k mindig első, a C-k pe-

¹⁴ PEÑA és kollégái kísérletsorozatának rövid elemző ismertetését magyarul lásd még FISER 2005: 33.

dig harmadik szótagokként jelennek meg. Ezért a teszt sztringjeit olyan szótagnégyesekre tagolva mutatták be, amelyeket a korábban használt szókból és nem-szókból állítottak elő úgy, hogy középső szótagjaikat kicserélték két-két új szótagra (vagyis ez esetben X^* -ként nem egy, hanem két szótag szerepelt). A kutatók úgy gondolták, ezek a vizsgálatok alátámasztották feltételezésüket: szerintük ugyanis az ingeranyagoknak ezekben az eloszlásaiban a statisztikai mutatók a nemszavaknak kedveztek, így a kísérleti személyeknek — ha az ilyen esetekben algebrai képletek helyett átmeneti valószínűségekre építenének — a mostani két tesztfázisban ezeket kellett volna grammatikusként megnevezniük.

Noha a gondolatmenet, ami ennek a két ellenőrző mérésnek a háttérében áll, önmagában véve meggyőzőnek tűnhet, a kísérletsorozat egésze felől nézve mégis logikai hibáról árulkodik. A bemutatott ingeranyagban ugyanis nemcsak a szünetek beiktatásának eredményeként tűntek fel olyan statisztikai mutatók, amelyek az AXC szavak felismerését támogatták, hanem már a kezdetektől fogva, a sztringeket ömlesztve bemutató változatban is jelen voltak.

SEIDENBERG, MACDONALD és SAFFRAN (2002) például kiszámolták, hogy abban a beszédmintában, amit PEÑÁék az öt főkísérlet során használtak, nemcsak az algebrai formulaként kezelt AXC előfordulásának (vagyis annak, hogy a szavak *puXki*, *beXga* és *tuXda* formájúak) volt 1,0 az esélye, de ezzel együtt még további öt tényezőnek is. Így annak, hogy a szavak első, valamint utolsó szótagjai felpattanó zárhanggal (azaz *p*-vel, *b*-vel és *t*-vel, illetve *k*-val, *g*-vel és *d*-vel) kezdődtek, hogy a két likvidát (*r*-et, *l*-et) és az egyetlen spiránst (*f*-et) tartalmazó szótagok csak *X* pozícióban jelenhettek meg, hogy a *pura* után mindig *ki* jött, és hogy a *pu*-t nem követhette *be*. Az AXC nyelv szavainak azonosítása szempontjából persze az sem mellékes, hogy 0,5 volt a valószínűsége a *be* előfordulásának *raki* és *ki* után, 0,33 pedig a *pu*-t követő *ra*-é.

Mindez korántsem elhanyagolható körülmény az adatok interpretációja szempontjából. Eléggyé nyilvánvaló, hogy a kísérleti eredményeket ezekre a statisztikai összefüggésekre alapozva éppúgy meg lehet magyarázni, mint a komplikáltabb, absztrakt szabályokra is építő kettős rendszer feltételezésével.¹⁵ Ily módon PEÑA és munkacsoportja — bár MARCUSék (1999) mérésének módszertani hiányosságaiból valóban sokat pótol — az eredeti célkitűzés egyik fontos elemét, a kétféle tanulás elkülönített tesztelését végül nem tudta megvalósítani.

Így van ez a vizsgálatsorozat másik apropójával, az egybevetetőséggel is. Vitathatatlan, hogy PEÑÁék öt kísérlete egymással összehasonlítható, de azokkal a korábbi tesztekkel, amelyekre reagálni kívánt, már csak korlátozottan. A SAFFRAN (1996) és MARCUS (1999) kutatócsoportjai által elvégzett vizsgálatok mind a résztvevők, mind pedig az ingeranyag tekintetében eléggé különböztek a PEÑÁékétól.

¹⁵ Vö. még FISER 2005: 33, TAUZIN 2010: 391.

Egyrészt a SAFFRAN- és MARCUS-féle csoport csecsemőinek, valamint PEÑÁék felnőttjeinek kognitív képességei (beleértve a tanulási stratégiákat is) közvetlenül aligha mérhetők össze. Másfelől az sem lényegtelen eltérés, hogy a három kutatócsoport egymáshoz képest heterogén komplexitású és időtartamú beszédmintákat állított össze. SAFFRANÉK négy, nyolcféle konzonánst és négyféle vokálist tartalmazó, három szótagos egyedi szót (tizenkét szótagtípust) mutattak be ömlesztve, két percen át. MARCUSÉK tizenhat, négyfajta mássalhangzóból és kétfajta magánhangzóból álló háromszavas ABA és ABB mondatot (nyolc szótagtípust) játszottak le egy másodperces (küszöb feletti) szünetekkel tagolva, ugyanennyi idő alatt. Mindeközben PEÑA és kollégái kilenc, kilencféle konzonánsból és ötféle vokálisból összetevődő három szótagos AXC szót (kilenc szótagtípust) hallgattattak meg tíz-tíz, majd harminc percen át folyamatosan, illetve tíz, valamint két percig, huszonöt milliszekundumos (küszöb alatti) szünetekkel elválasztva.

16. SAFFRAN (1996), MARCUS (1999) és PEÑA (2002) kutatócsoportjainak mérései a kognitív pszichológiában olyan szakmai lavinát indítottak el, amely újabb és újabb, az előzőekhez képest valamilyen szempontból módosított kísérletek és metodológiai megjegyzések sorát hozta magával.¹⁶ Bár ezek sok, főként módszertani részletkérdést tisztáztak, a vitát lényegileg, vagyis konceptuálisan nem igazán vitték előre.

Egyrészt a két tábor közti véleménykülönbség jelenleg is megvan, és az álláspontok manapság is éppen ugyanott és ugyanúgy térnek el egymástól, mint korábban. Másfelől a további vizsgálatok valamennyi verziója azzal a tanulsággal szolgál, amit már a fentiekben ismertetett három híres mérőcsoport eredményeiből is levonhattunk. Az algebrai műveletek pártján állók úgy igyekeztek ennek a tanulási formának a létezését alátámasztani, hogy mindeközben kísérleteik feladatai átmenet-valószínűségekre alapozva is megoldhatók voltak. A statisztikai tanulás képviselői viszont tesztjeikben olyan hangsorokkal szembesítették alanyaikat, amelyek elemei között csak együttjárási gyakorisági összefüggések voltak, logikai jellegűek nem.

Nincs tehát olyan kísérlet, amely a nyelvtannak akár a statisztikai elsajátítását, akár az algebrai megismerését kétséget kizáróan cáfolná, de olyan sincs, amely egy statisztikai korrelációktól mentes, tisztán logikai nyelv grammatikai struktúrájának felfedezését szemléltetné. Ugyanakkor van példa arra, hogy az emberek megtanuljanak egy olyan nyelvtant, amely egyedül az elemek közti átmenet-valószínűségekkal jellemezhető. Ez már önmagában véve elégséges ahhoz,

¹⁶ Ízelítőül lásd csak a MARCUS és társai (1999) kiváltotta legelső reakciókat a nemzetközi szakirodalomból: ALTMANN–DIENES (1999), CHRISTIANSEN–CURTIN (1999), EIMAS (1999), MARCUS (1999a, 1999b, 1999c, 1999d, 1999e), MCCLELLAND–PLAUT (1999), NEGISHI (1999), SEIDENBERG–ELMAN (1999a, 1999b).

hogy a statisztikai tanulás lehetőségével mindenképpen, az algebrai stratégia eshetőségével pedig inkább csak feltételesen számoljunk. Mégis akadnak, akik a statisztikai folyamatokat a kísérleti ingeranyagok nyelvtani tulajdonságai láttán sem tartják valószínűbbnek az algebrai műveleteknél.

17. Némelyek például egyenesen úgy gondolják, az alanyok túl jól teljesítettek ahhoz, hogy statisztikai úton tanulhassanak. A MARCUSÉK-féle mérésben (1999) a hét hónapos csecsemőknek nem a hallott ingerek, hanem azoknak általánosabb, repetitív jegyei között kellett együttjárás gyakorisági összefüggéseket észrevenniük ahhoz, hogy a bemutatott szótagkombinációkat ABA és ABB típusokként kezelhessék, majd korábban nem tapasztalt elemekre is alkalmazhassák. PEÑA és munkatársai harmadik és ötödik tesztjében (2002) pedig a felnőtteknek az AXC szekvenciák szavai között nem is közvetlen (szomszédos szótagok közötti), hanem közvetett (a nem szomszédos A_i-k és C_i-k közötti) átmenet-valószínűségeket kellett felfedezniük ahhoz, hogy ezt a mintázatot ismeretlen hangkapcsolatokra is kiterjeszthessék.

Azok, akik ilyen érvekkel kétségbe vonják a nyelvtani folyamatok statisztikai jellegét, a természetes nyelvek megismerésének menetéből vett tapasztalatok alapján tulajdonképpen azt tartják kérdésesnek, hogy olyan hanganyagok átmenet-valószínűségeit lehetséges detektálni, illetve akár soha nem hallott szótagokra vonatkoztatni, amelyekben a bemutatott típusok számához és/vagy statisztikai bonyolultságához képest viszonylag kevés a tokenek mennyisége. Lényegében erre hivatkozik például ENDRESS és a PEÑA-féle cikk (2002) egyik társszerzője, BONATTI is annak a kísérletsorozatuknak (ENDRESS–BONATTI 2007) az elemzése során, amelyben a tréningiszöveg komplexitását a szoktatási idő függvényében variálták. Mérési adataikat értékelve arra a következtetésre jutottak, hogy statisztikai tanulás csak megfelelő mennyiségű mintavétel esetében célravezető, egyébként az elemek együttjárás gyakorisági jellemzőinél a köztük lévő általános logikai összefüggések hatékonyabb grammatikai markereknek látszanak.

Nem árt azonban szem előtt tartani, hogy a vizsgálati ingeranyag a természetes beszéd hangsoraihoz képest kivétel nélkül lényegesen egyszerűbb volt, ami pedig pozitívan befolyásolta a résztvevők teljesítményét. Mindenekelőtt az emberi nyelvekben a kísérleti példakénál jóval bonyolultabb fonotaktikai rendezettségű hangsorok fordulnak elő. Emellett a tanító szövegeket mindig géppel hozták létre úgy, hogy elemeiket prozódiai, intonációs és hangsúlyozási jegyeik tekintetében homogenizálták, ezzel is csökkentve a beszédminta akusztikai komplexitását. A tanulás hatékonyságát ezen túl tovább növeli az a körülmény, hogyha a kísérletben használt műnyelv hangsorait ráadásul tagolt hármasokban játsszák le: az AXC szekvenciákkal kapcsolatos PEÑA-féle vizsgálatokban (2002) láttuk, hogy az alanyok feladat-megoldási sikeressége pusztán a szünetek beiktatása miatt jelentősen megváltozott. Ezek tükrében pedig nem meglepő, hogy a mesterséges és a természetes ingeranyaggal végzett mérések nem

mindig zárulnak ugyanazzal az eredménnyel. Az egy korosztályba tartozó anyok a műnyelvet használó tesztekben gyakran teljesítenek jobban, mint anyanyelvükön vagy akár valamely más természetes, de számukra idegen nyelven.

Az várható volt, hogy MARCUSék (1999) hét hónapos csecsemői képesek legyenek az ABA és az ABB szerkezeteket differenciálni, hiszen nyelvelsajátítási kísérletekből tudjuk, hogy a jóval bonyolultabb anyanyelvi szótagkombinációk memorizálása is éppen erre az időszakra (a hatodik és tizenkettedik hónapok közé) tehető.¹⁷ Az a mérési eredmény viszont, amit GERVAIN és kollégái (2008) azon neurolingvisztikai kísérletük során kaptak, amelyben a MARCUSék által használt szekvenciákat újszülötteknek mutatták be, csak akkor érthető meg minden vonatkozásában, ha figyelembe vesszük, hogy ez esetben a természetes nyelvekben előfordulókhöz képest egyszerűbb, mesterséges nyelvi ingerekről van szó. A jobb és bal temporális, illetve a bal frontális területeknek az ismétlődő szótagokra adott megnövekedett válaszaiból ugyanis az látszott, hogy nemcsak a hét hónaposok, de már a néhány napos csecsemők is el tudják különíteni¹⁸ a közvetlen ismétlődést tartalmazó (*mubaba, talulu* stb.) elemsorokat a kontrollként használt nem redundatív (*mubage, talupi* stb.) szekvenciáktól.¹⁹

¹⁷ Ez idő tájt a babák reakciói világosan eltérnek egymástól, amikor például anyanyelvük szavait jól formált hangsorokként és visszafelé is lejátszzuk nekik. Ez pedig a fonotaktikai mintázatok elsajátítására utal: a kétféle hanganyagának az akusztikai komplexitása egymással teljesen megegyezik, ez utóbbira alapozva tehát nem lehetne különbséget tenni az anyanyelvi és az ellentétes hangsorrendű szavak között (vö. még CSIBRA 2003: 258).

¹⁸ Mivel itt újszülöttek habituáció nélküli teljesítményéről van szó, felvetődhet annak a lehetősége, hogy náluk ne tanulást, hanem valamiféle „veleszületett” tudást feltételezzünk. Ez azonban kérdésünk, a statisztikai-algebrai tanulás vitája szempontjából semmin nem változtat. Egyrészt ekkor arról lenne szó, hogy a MARCUS-féle tesztek (1999) valójában nem is a tanulást, hanem a tudást mérték. Másrészt pedig — ahogy GERVAINék második kísérletéből (2008), valamint PEÑA és kutatócsoportja vizsgálataiból (2002) kitűnik — az elkülönítés korlátozott, a közvetett ismétlődést tartalmazó szerkezetekre nem terjed ki, így mindenképpen szükség van arra, hogy további tanulás útján ez a képesség kiteljesedjen.

¹⁹ A módszertani alaposág kedvéért fontos megemlíteni, hogy GERVAIN és munkatársai kísérletükhöz egy olyan, még újnak mondható műszert használtak, amelynek segítségével gyenge infravörös fényt sugároznak be a koponyán keresztül az agyba, és annak visszaverődéséből következtetnek az érintett terület oxigénfogyasztására (vö. CSIBRA 2003: 261). Azonban az optikai topográfának (az ún. „NIRS”-nek, vagyis a „Near Infrared Spectroscopy”-nak) — csakúgy, mint minden neurolingvisztikai mérőeszköznek — az értékei (különösen ilyen életkorban és az ingerváltáshoz kapcsolt újdonságmarkerekre alapozva) eléggé széles skálán interpretálhatók. Így a vizsgálatból származó adatok önmagukban nem, csak a többi, viselkedéses reakcióra alapozó kísérlet tükrében értelmezhetők.

Azóta persze ezt a GERVAIN-féle tesztet (2008) számos olyan méréssel ismételték meg, amelyek a szopásgyakorlás megnövekedésére mint újdonságmarkerre (tehát az ún. „HAS”, vagyis a „High Amplitude Sucking” procedúrára) alapozva megbízhatóbb eredményeket szolgáltatnak (ezekről részletesebben lásd GERVAIN és mtsai 2011: 40–41). Annak, hogy itt mégis ezt a neurolingvisztikai kísérletet és nem a kognitív pszichológiai ellenőrző változataikat idézem, oka van. Azon túl, hogy a következtetéseket nem befolyásolja, hiszen a viselkedéses vizsgálatok az eredeti teszt eredményeit megerősítették, GERVAINék munkájának a statisztikai-algebrai tanulás vitája szempontjából nem elhanyagolható tudománytörténeti jelentősége van. Míg korábban e kérdéskör-

A művi és a természetes ingeranyag eltéréseiből adódhat az a látszólagos el-
lentmondás is, amit GOMEZ és GERKEN mesterséges mondattani tesztjének (1999),
illetve a SANTELMANN–JUSZYK páros anyanyelvi szintaktikai kísérletének (1998)
eredményei között tapasztalhatunk.

GOMEZ és GERKEN egy véges állapotú automatával modellezhető, nyolcsza-
vas szótárral dolgozó grammatika segítségével generált különböző bonyolult-
ságú mondatokat (például *vot pel jic, pel tam pel jic, pel tam jic rud tam rud*).
Egyéves csecsemőket kevesebb mint két percig szoktattak az ezekből véletlen-
szerűen összeállított mintához. Ezt követően a tesztfázisban agrammatikus szek-
venciákat (*tam jic rud vot, tam tam jic rud pel, jic pel jic rud tam vot* stb.), illetve
olyan sztringeket mutattak nekik, amelyeket a habituációs szakaszban nem hall-
hattak, de az automata elfogadta őket mint jól formált mondatokat (*vot jic rud
tam, pel tam jic rud tam, vot pel jic rud tam jic* stb.). A mérés során a kutatók a
klasszikus gyakorlatnak megfelelően az ingerpreferencia-jelzőkre alapoztak, és a
fejfordításos paradigmát használták. Azt találták, hogy a babák minden esetben a
grammatikus mondatokra figyeltek inkább,²⁰ még akkor is, ha a teszt során a
nyelvtanilag jól formált mondatokat egy új szótárt használva hozták létre.

Míg ebben a kísérletsorozatban az egyévesek képesek voltak akár hattagú
szintaktikai formulák között differenciálni, SANTELMANN és JUSZYK (1998)
empirikus eredményei arra utalnak, hogy az angol folyamatos jelen idejű monda-
tok grammatikai markereinek összetartozását a babák legkorábban másfél éves
koruktól észlelik. E kísérlet során a kutatók tizenöt és tizennyolc hónaposoknak
habituáció nélkül, egy rövid gyakorlás után egy grammatikus és agrammatikus
mondatokat tartalmazó hanganyagot mutattak be. A lejátszott szövegben így két-
féle nyelvtani szerkezet fordult elő egyenlő arányban: egyes szám harmadik sze-
mélyű segédigés folyamatos jelen idejű (*is ... -ing*) grammatikus mondatokkal
(*At the bakery, everybody is baking bread. One person is mixing the flour and
water together. Somebody else is adding salt and yeast.* stb.) keverve olyan nem
jól formált mondatok, amelyeket az előbbieik módosításával hoztak létre úgy,
hogy a segédige típusát egy nem odaillőre (*can ... -ing*) cserélték (*At the bakery,*

ben teszteket főleg nagyobb babákon és felnőtteken végeztek, ebben a vizsgálatban GERVAIN és
kutatócsoportja a MARCUSÉK (1999) mérését újszülötteken kívánta megismételni — nem véletlen,
hogy a kísérletet vezető kutatók között ott találjuk PEÑÁT és egyik társszerzőjét, MEHLERT is (PEÑA
és mtsai 2002).

²⁰ Az, hogy a csecsemők az ismert vagy az új ingereket preferálják-e inkább, egyébként nem egé-
szen egyértelmű, ám ez a mérés validitásán szerencsére mit sem változtat. Míg például SAFFRANÉK
(1996) és MARCUSÉK (1999) nyolc, illetve hét hónaposai az ismeretlen sztringeket részesítették
előnyben, GOMEZ és GERKEN (1999) tesztjében — csakúgy, mint az összehasonlításképp ezután be-
mutatandó SANTELMANN–JUSZYK-féle mérésben (1998) — fordítva, holott az életkori jellemzők
alapján az egyéves babáknak szintén az újdonságokat kellene kitüntetniük figyelmükkel. De akár-
melyiket is kísérik nagyobb érdeklődéssel a gyerekek, bármilyen ingertípus preferenciája csak ak-
kor lehetséges, ha azt a többitől elkülönítik — az efféle tesztek pedig a babáknak éppen erre a
diszkriminálási képességére kérdeznék rá. (Mindezt vö. még LUKÁCS 2005: 39.)

everybody can baking bread. One person can mixing the flour and water together. Somebody else can adding salt and yeast. stb.). A kísérletvezetők a teljesítmény mérésére ez esetben is a szokásos fejfordításos preferencia paradigmát használták, és azt látták, hogy míg a tizenöt hónaposok nem különböztették meg a kétféle szintaktikai struktúrát, a másfél évesek szignifikánsan tovább figyeltek a grammatikus példákra, legalábbis akkor, ha a két nyelvtani marker közötti távolság nem volt több három szótagnyinál (például *At the bakery, everybody is often baking bread.*). Ennél messzebbre nyúló átmenet-valószínűségek esetében ugyanis reakciójuk a fiatalabbakéval egyezett meg (például *At the bakery, everybody is cheerfully baking bread.*).²¹

18. Úgy tűnik, a mesterséges és a természetes ingeranyaggal végzett kísérletek eredményei összeegyeztethetők, maguk a feladatok pedig összességében inkább a grammatika statisztikai elsajátítását valószínűsítik, igaz, az algebrai nyelvtanulást sem zárják ki teljesen. Ha mindezek mellé számításba vesszünk még két további vizsgálati tényezőt is, a grammatikai műveletek statisztikai jellegére újabb megerősítő érveket találhatunk.

A megismerés graduális, átmenet-valószínűségekre alapozó jellege jól látszik mindenekelőtt abból, hogy a kísérleti személyek teljesítménye és az ingeranyag statisztikai komplexitása összefüggést mutatott. Az alanyok számára a kevésbé bonyolult kapcsolódási minták memorizálása egyszerűbb volt, mint az összetettebb szerkezetek felismerése.

GERVAINÉK (2008) méréseik során az újszülötteknél az ABC kontrollstruktúrához (*penaku* stb.) képest csak a percepciósan feltűnőbb, közvetlen reduplikációt tartalmazó ABB szekvenciák (*penana* stb.) észlelésekor találtak neurális reakcióbeli eltérést, az indirekt ismétlődést mutató ABA formulák (*napena* stb.) kapcsán nem. A MARCUS- (1999), illetve a PEÑA-munkacsoportok (2002) tesztjeit együttesen szemlélve pedig az látszik, hogy a közvetlen szomszédsági átmenet-valószínűségek, így például az ABB reduplikatív elemeinek felfedezése könnyebb az olyan közvetett statisztikai összefüggések detektálásánál, mint amilyen mondjuk az AXC típusé. Mindez egyébiránt teljesen összevág ENDRESS, SCHOLL és MEHLER (2005) vizsgálataival is, akik a MARCUS-féle szerkezetekhez hasonló, de hét szótagú szekvenciákat (például ABCDDEF) alkalmazva azt tapasztalták, hogy a francia felnőtt kísérleti személyek csak akkor tanulták meg az ismétlődő elemekből felépülő struktúrákat, ha azok egy percepciósan feltűnőbb pozícióban, a mondat valamelyik szélén helyezkedtek el (például ABCDEFF).

A tanulás átmenet-valószínűségi alapjait támogatják az elvégzett vizsgálatoknak a szoktatás mértéke és a teljesítmény kapcsolatára utaló momentumai is. Azt, hogy a habituáció időtartama és az eredményesség között kor-

²¹ A kísérlet rövid, magyar nyelvű bemutatását lásd például LUKÁCS 2005: 40, emellett vö. továbbá CRYSTAL 2003: 304.

reláció van, természetesen nemcsak a statisztikai tanulás során kell feltennünk, de az algebrai megismerés esetén is. Ami az utóbbival szemben az előbbi mechanizmusra enged következtetni, az nem a megfelelés megléte, hanem a jellege: annak a függvénynek a mintázata, amit a két mutatónak a viszonya kirajzol.

Az algebrai tanulás tulajdonképpen egy heurisztikus művelet. Így a kísérleti személyeknek az olyan tesztekben, amelyek a szoktatás során nem tapasztalt tokeneket tartalmaznak, mindaddig, míg nem fedezték fel az adott nyelv grammatikáját jellemző általános formulát, nem kellene semmiféle teljesítményt mutatniuk (a résztvevők válaszadási százaléktételeinek egyenként véletlenszerűsége kellene tükrözniük). Ám ha a habituáció idejét úgy megnövelik, hogy az alanyok elérjék a szabályfelismeréssel járó heurékaélményt, ez után a revelatív pont után ezeknek a feladatoknak a megoldásában hirtelen mindent tudniuk kellene (a nem véletlenszerűen válaszoló résztvevők grammatikai döntéseinek egytől egyig helyesnek kellene lenniük). A kísérleti eredmények azonban soha nem ilyen szélsőségesek: a helyett a két eltérő magasságú konstans szakaszból álló lépcsős függvény helyett, amit az algebrai megismerés esetében látnunk kellene, valamilyen más, jellegzetesen íves tanulásmenetet tapasztalhatunk.

Persze a logikai mechanizmus mellett állást fogláló kutatók gyakran élnek azazal a kritikai észrevétellel, miszerint a mérési adatok éppen a statisztikai folyamatot nem támogathatják, hiszen ha mégis ezt tennék, akkor a habituáció mértékének és a tanulás eredményességének egyenes arányosságban kellene állnia egymással. Ez az érv, amely a statisztikai tanulás e két mutatója esetében egy lineáris függvényt feltételez, azonban nem vesz számításba minden tényezőt. Az eredményesség ugyanis még kísérleti helyzetben sem csak a szoktatás intervallumától függ, hanem az ingeranyag komplexitásától is, a beszédminta grammatikájának bonyolultsága pedig nem egyenesen arányos a megváltoztatott elemek számával. Két szótag cseréjének hatása a nyelvtan összetettségére nem kétszerese annak, mint amit egy szótag módosításával kapnánk, hanem az adott elem átmenetvalószínűség-értékeitől függően ennél nagyobb vagy kisebb. Ha mindez még ráadásul a szoktatási idő paramétereivel is kombinálódik, a tanulás menete a habituáció hosszának és a teljesítménynek az összefüggése tekintetében aligha lehet egyenes vonalú és egyenletes. Egy korpusz két átmenet-valószínűségi korrelációjának felfedezése sem pontosan kétszeres tudással jár ahhoz képest, amit egy együttjárási gyakorisági mutató detektálása hoz magával — az előbbieken túl már csak azért sem, mert itt nem abszolút, hanem relatív (egymáshoz, az addig hallott elemekhez együttesen viszonyított) eloszlásokról van szó.

Jó példa az adatok ilyen szempontból eltérő interpretációira többek között PEÑÁÉK (2002) negyedik tesztjének elemzése. Ez az a vizsgálat, amelynek során hosszabb familiarizációs szakaszt követően figyelték meg az alanyok reakcióit, a résztvevők pedig szignifikánsan többször választották az ingeranyag azon elem-sorait, amelyek nem feleltek meg az AXC szekvenciának, ha a szótaghármások-

ba korábban X pozícióban nem tapasztalt elemeket vegyítették. A kísérletvezetők a nem szó szerinti elemek kiváltotta figyelmet úgy értelmezték, hogy a megnövelt tréningfázis következtében a memórianyomok erősebbek lettek, míg a szerkezetre vonatkozó általános jegyek eltompultak. Ennél azonban — ahogy FISER is írja — „egyszerűbbnek tűnik az a magyarázat, hogy a megnyújtott szoktatási idő megerősítette a struktúrára vonatkozó információk statisztikus kiszűrését, és ennek megfelelően, amikor a teszt során a struktúrát elrontó elemek kerültek a »szabály-szó« szótaghármába, az alanyok rögtön észlelték azt” (2005: 34).

19. Ahhoz, hogy a statisztikai tanulás mellett az algebrai, illetve a kétféle (hol elvont kategóriák közötti logikai megfelelésekre, hol pedig empirikus adatok előfordulási rendjére alapozó) grammatikai műveletek gondolatát komolyan számításba lehessen venni, olyan tesztekre van szükség, amelyek (egyes) feladatait semmiképpen nem lehet együttjárási frekventatív összefüggésekre támaszkodva megoldani (FISER 2005: 34–35).²² Az viszont, hogy a gyakorlatban tervezhető-e egyáltalán ilyen kísérletek, már fölöttébb kétséges, az általános szabályok kiemelésének a statisztikai megismeréstől való elkülönítése ugyanis már teoretikus értelemben is megoldhatatlannak látszik.²³

Absztrakt grammatikai változókkal csak akkor tudunk operálni (legyen ez akár szerzett, akár innát képességünk), ha ezeket egy ingerkészletből korábban már felfedeztük, azaz ha előzőleg nagyszámú egyszerűbb és komplexebb valószínűségi korrelációt ismertünk meg a hangsorok egyes konkrét részletei között (FISER 2005: 33). Ebből a szemszögből nézve tehát a „statisztikai tanulás” terminus egy olyan tautologikus kifejezés (vö. FISER 2005: 27), ami létét alapvetően a tudománytörténetének köszönheti: annak, hogy rivális elméletként, a kognitív pszicholingvisztikát a kezdetektől meghatározó szabályalapú koncepció után jelent meg.

Mivel pedig ily módon a *t a n u l á s* csak egyféle — méghozzá: statisztikai — lehet (FISER 2005: 27), a folyamat jellegét firtató kérdés tulajdonképpen irreális, vita legfeljebb az ebből származó *t u d á s* természetéről folyhat. Ez a felvetés azonban csak akkor értelmes, ha lehetséges az, hogy a grammatika elsajátítása statisztikai alapú legyen, ám az ennek eredményeként kibontakozó tudás (egészében vagy részben) algebrai.

20. Ha a tanulást és a tudást a nyelvten kapcsán úgy választjuk ketté, hogy az előbbit statisztikainak, az utóbbit pedig algebrai tekintjük, fennáll a veszélye, hogy ezzel észrevétlenül magát a „statisztikai tanulás” fogalmát is radikálisan *á t é r t e l m e z z ü k*. Mivel ez esetben a tanulás és a tudás minőségileg különböznek egymástól, könnyen az az érzésünk támadhat, hogy az előbbi csak egy olyan előkészítése az utóbbinak, ami a nyelvelsajátító számára különösebb kognitív kihívást nem jelenthet. Egy egyszerű hangsor-memorizálási folyamat,

²² További adalékként TAUZIN 2010: 391, 397.

²³ Ehhez lásd még LUKÁCS hasonló tartalmú, ám kevésbé határozott kijelentéseit (2005: 39–40).

ami pusztán egy nyelvtani értelemben strukturálatlan adatfelhalmozódással jár. E megközelítésben tehát a „statisztikai tanulás” kifejezés szemantikája drámaian *l e e g y s z e r ű s ö d i k*: a megismerés menete tulajdonképpen nyelvi elemek listázásává redukálódik, és csak az általánosított algebrai formulák minősülnek valódi szerkezeti információkat hordozó, grammatikai tényezőknek. Eszerint az átmenet-valószínűségek mindössze közvetlen szomszédossági viszonyokat jelelhetnek (távolabbi elemek közti relációkat nem jelezhetnek), és nem terjeszthetők ki új elemekre (azaz nem produktívak).

A statisztikai tanulásnak ez a felfogása korántsem példa nélküli a szakirodalomban. Tulajdonképpen ilyen a nyelvészetben CHOMSKY elképzelése is. Ez nem is túl meglepő: jól tudjuk, hogy éppen ez az a nyelvtudományi iskola, amire hagyatkozva a kognitív pszichológiában a mentális nyelvtant algebrai jellegűnek tételezték fel.

21. CHOMSKY az azóta mind a nyelvészetben, mind pedig a kognitív pszichológiában klasszikussá vált példapárral, a *Colorless green ideas sleep furiously* és a *Furiously sleep ideas green colorless* hipotetikus mondatokkal²⁴ azt igyekszik szemléltetni, hogy miután bármelyik angol anyanyelvű felismeri, hogy csak az első a grammatikus (1957/2003: 17), a „beszélőnek az a képessége, hogy nyelvtanilag helyes megnyilatkozásokat hoz létre és ismer fel, nyilvánvalóan nem olyan fogalmakon alapszik[,] mint a statisztikai gyakoriság és hasonlók” (1957/2003: 19). Szerinte ugyanis az, hogy az előbbi két mondat közül egyik sem hangzott el soha (a szavak közti átmenet-valószínűség 0,0), de az emberek mégis tudják, hogy grammatikailag melyik a jól formált, annak az egyértelmű jele, hogy a nyelvtani tudás nem alapulhat a nyelvi elemek semmiféle olyan mutatóján, ami az előfordulásukra vonatkozik (vö. még 1957/2003: 20).²⁵

Jól látszik, hogy ez a vélekedés a statisztikai tanulás kapcsán az átmenet-valószínűségekkel mint grammatikai strukturális tényezőkkel egyáltalán nem számol, így azzal sem, hogy a megtapasztalt együjtjárás gyakorisági mintázatok egymástól távolabb elhelyezkedő elemekre is vonatkozhatnak, és bármiféle produktív nyelvtani művelet szerkezeti alapjaiként szolgálhatnak. Ez a felfogás azonban semmivel nem indokolható, hiszen a statisztikai megismerés épp annyira grammatikai természetű tudást eredményez, mint az algebrai módszer.

Eléggő nyilvánvaló, hogy ha átmenet-valószínűségeket tanulunk, akkor sem csupán egy egyszerű, közvetlen szomszédossági viszonyokon alapuló adathalmazt rögzítünk listaszerűen elménkben, hanem ugyanúgy a nyelvi elemek struktúrájáról kapunk információkat, mint az algebrai út esetén. Ami kétségtelenül más, az az algebrai és a statisztikai grammatika szervező elve: az előbbi kétértékű

²⁴ A mondatok jelentése ezúttal nem mellékes, így fordításaikat is közlöm. Az előbbi a *Szintelen zöld eszmék alszanak dühösen* formában hangzana magyarul, az utóbbi pedig valami olyasmi lenne, mint hogy *Dühösen alszanak eszmék zöld szintelen* (vö. CHOMSKY 1957/2003: 17).

²⁵ Vö. még PINKER 1994/2006: 86, 91–92, GERVAIN 2014: 128, PLÉH 2014: 53–54.

(bináris), míg az utóbbi skálaszerű (skaláris). Ez a gyakorlatban azt jelenti, hogy mindkét rendszer működése *p r o d u k t í v*, ám meglehetősen eltérő módon.

Valahogy úgy, ahogy a kétféle stratégia különbözik akkor, amikor ezeket manapság a számítógépes nyelvészeten mint nyelvtechnológiai módszert használják, hogy a gépeket természetes nyelvekkel kapcsolatos feladatok elvégzésére tegyék alkalmassá. „A szabályok tipikusan »mindent vagy semmit« alapon működnek, tehát vagy érvényesül egy szabály a feldolgozás során, vagy nem érvényesül, közbülső állapot nem létezik. Ezzel szembeállítható az újabban elterjedt statisztikai módszer, ahol a rendszer természetes nyelvi szövegekből [...] automatikusan von ki mintákat és ezeket alkalmazza más nyelvi megnyilvánulások feldolgozására. A mintákhoz tipikusan valószínűségeket rendel az algoritmus és újabb szövegek feldolgozásakor vagy generálásakor arra törekszik, hogy az együttes valószínűség maximalizálása mellett valósítsa meg a mintaillesztést” (BABARCZY–GÁBOR–SERÉNY 2014: 414).

Egy algebrai nyelvtan egy adott formulája tehát akkor és csak akkor alkalmazható egy még soha nem hallott elemre, ha az új elem vele strukturálisan teljesen *a z o n o s* — ez az az eset, amit a nyelvtudományban úgy fogalmazznak meg, hogy a *s z a b á l y* vonatkozik az adott elemre. Ezzel szemben egy statisztikai grammatika egy bizonyos valószínűségi mintázata akkor terjeszthető ki egy korábban elő nem forduló elemre, ha annak felépítése *h a s o n l í t* rá, még hozzá a beszélő-hallgató által már ismert szerkezetek közül az adott együtjárás gyakorisági szekvenciához hasonlít leginkább — ez az, ami a nyelvészeten az *a n a l ó g i a* korszerű értelmezése lehetne.

CHOMSKY híres példáiban az angol anyanyelvűek tehát először is szemantikai jegyeik miatt feltehetőleg mindkét mondatot szokatlannak tartanak, ítéleteik pedig csak a furcsaság mértékében térnének el: az egyiket bizonyos speciális helyzetekben (mondjuk metaforikusan) el tudnák képzelni, a másik megjelenését viszont aligha. Másodszor pedig nem azért tudnák, hogy a két mondat közül melyik a kevésbé szokatlan felépítésű (*Colorless green ideas sleep furiously*), és melyik a nagyon különös (*Furiously sleep ideas green colorless*), mert az egyiket már hallották, a másikat meg nem, hanem azért, mert olyanokat már sokszor hallottak, amelyek az egyikhez strukturálisan hasonlítanak (*Hopeless sad guys sleep deeply* stb.), de olyanokat nemigen, amelyek a másikhöz szerkezetileg hasonlítanak (**Deeply sleep guys sad hopeless* stb.).

22. A kognitív pszichológiában megfigyelhető, hogy az algebrai formulák oldalán állók gyakran érvelnek úgy, hogy a statisztikai tanulás legfeljebb a szóhangalakok egyszerű memorizálásában lehet hatékony, a grammatikának a nagyobb kognitív teljesítményt igénylő felfedezésére viszont nem alkalmas.²⁶ Ez-

²⁶ E tekintetben igen jellegzetes LUKÁCS és társszerzői kijelentése a legújabb pszicholingvisztikai kézikönyvben: „SAFFRAN és munkatársai azért választották a szószzerű egységek szegmentálását, mert a nyelvelsajátításnak ezt a területét feltehetőleg mindenki tapasztalatvezéreltnek [értsd:

zel együtt az átmenet-valószínűségeken alapuló megismerést nem tartják pusztán strukturálatlan adatfelhalmozódásnak, ám egymástól távolabb álló elemek közti viszonyokra vonatkoztathatóknak és produktívnak sem.

MARCUSÉK (1999) és PEÑÁÉK (2002) a kísérleteiket tulajdonképpen arra a gondolatra alapozták, hogy az együttjárási gyakorisági mintázatok nem terjeszthetők ki új nyelvi elemekre, de szótanulásra megfelelnek, ahogy ezt az utóbbit SAFFRAN és csoportja (1996) tesztjéből is láthattuk. Emlékezzünk, hogy a mérések protokolljának megtervezésekor mind MARCUS, mind PEÑA kutatócsoportja abból az előfeltevésből indult ki, hogy ha a feladatokban az ABA és ABB, valamint az AXC formulákat olyan tokenekkel tesztelik, amelyek a habituáció idején nem hallhatók, és az alanyok ennek ellenére jól teljesítenek, akkor történt tanulás, de biztosan nem statisztikai úton. Ez utóbbihoz tehát szerintük tokenszintű tapasztalatra, vagyis az adott elemek korábbi előfordulására lenne szükség.

Ily módon az algebrai stratégia képviselői a „statisztikai tanulás” fogalmának egyszerűen más jeleket tulajdonítanak anélkül, hogy ennek tudatában volnának. A statisztikai megismerés oldalán állók a valószínűségi mintázatok el-sajátítását egy olyan grammatikai strukturák megismerésével járó folyamatként értelmezik, amely produktívan működtethető. Ezzel szemben az algebrai tanulás mellett érvelők ugyanezt egy improduktív rendszernek gondolják. Annak a pátthelyzetnek, ami az algebrai–statisztikai tanulás dilemmájában előállt, az alapvető oka az, hogy a vita résztvevői valójában nem is ugyanarról a statisztikai tanulásról beszélnek — a két malomban őrlés pedig az érdemi továbblépést erősen megnehezíti.²⁷

Ebből a szemszögből nézve szinte várható, hogy az algebrai stratégia képviselői még mindig azon az állásponton legyenek, amit CHOMSKY a vita kirobbanása előtt, az ötvenes években fogalmazott meg: „Törekedni lehetne egy összetettebb kapcsolat kialakítására a statisztikai eloszlás és a mondattani szerkezet között, mint a statisztikai gyakoriság fentebb elutasított egyszerű modellje. Semmiképpen sem szeretnék amellett érvelni, hogy az ilyen kapcsolat elképzelhetetlen, nem tudok azonban e tekintetben egyetlen olyan javaslatról sem, amelyeknek ne lennének nyilvánvaló fogyatékoságai” (1957/2003: 20).

23. A statisztikai tanulás leegyszerűsített felfogása a nyelvészetben is érezteti hatását, ami persze áttételesen a kognitív pszichológia lingvisztikai elméleteibe is bekerül. Ennek köszönhető, hogy e tudományterületek grammatikai leírásaiban mint módszertani elv tovább erősödik két, korábban is eléggé markánsan jelen

statisztikainak] gondolja. [...] A szavak sorrendjét szervező szabályszerűségeknek (a szintaxisnak) a helyzete ugyanakkor kevésbé egyértelmű” (2014: 511)

²⁷ Lényegében a „statisztikai tanulás” terminusnak erre a kétféle szemantikájára hívja fel a figyelmet PEÑÁÉK kísérleti eredményeivel (2002) vitatkozva, CHOMSKY iskolapéldává lett két mondatának kritikai elemzése kapcsán SEIDENBERG, MACDONALD és SAFFRAN (2002) is. Mindezt vö. még FISER 2005: 25–26, 34.

lévő s z e m b e á l l í t ó p á r (dichotómia): a nyelvtan és a lexikon, illetve ezzel szoros összefüggésben a szabály és az analógia ellentételező szétválasztása.

Abból adódóan, hogy CHOMSKY a statisztikai tanulást pusztán memorizálási folyamatnak tartja, ami a szabályalapú nyelvtan felől nézve egy strukturálatlan adathalmazt eredményez, felfogásában a szótár felépítése és a grammatika működése teljesen külön utakon halad. A nyelvtudomány klasszikus nyelvtanai a közösségi (szociális) nyelv történetének vonatkozásában már ezt megelőzően is így kezelték a grammatikai és a lexikális kérdéseket,²⁸ most ez az elgondolás megjelenik a nyelvészetben, majd ennek nyomán a kognitív pszichológiában az egyes személyek (mentális) nyelvi fejlődése kapcsán is. A nyelv szociális és mentális szerveződési szintjét vizsgáló nyelvtudományi kutatások, valamint a nyelvészeti és kognitív pszichológiai diszciplínák pedig — körkörös hivatkozásokkal — ezután a nyelvtani és a lexikális jellemzők elkülönítését illetően egymás érveit látszanak megtámogatni.

De annak is meglesz a következménye mind a nyelvtudományban, mind a pszichológiában, hogy CHOMSKY grammatikaelméletében kizárólag csak az algebrai formuláknak megfeleltethető szerkezetekkel, vagyis a szabályokkal számol. E tekintetben olyan szigorú, hogy a valószínűségi mintázatokhoz igazodó analógia fogalmát — egy tisztán logikai nyelvtani leírás megteremtése érdekében — egyenesen módszertani tiltólistára helyezi (1986).²⁹ Ezzel viszont még inkább meggyengíti ennek a hasonlósági alapú magyarázó elvnek az egyébként sem túl erős státusát a nyelvészetben, illetve ezen keresztül a kognitív tudományokban.³⁰

24. A fentieknek megfelelően CHOMSKY elkülöníti a veleszületett n y e l v - t a n i t u d á s t a későbbi l e x i k a i t a n u l á s i folyamatoktól, és a

²⁸ A történeti-összehasonlító nyelvészet álláspontja szerint míg a szókészletben a nyelvi érintkezések eredményeként nagy számmal megjelenhetnek más nyelvekből származó átvételek, a grammatika ilyen szempontból sokkal korlátozottabban változhat. Az alapnyelv nyelvtanának hagyományozódása szigorúan egyágú. Míg tehát a lexikális elemek elegyedése a nyelvek között elképzelhető, a nyelvtan esetében ugyanez aligha lehetséges (RÓNA-TAS 1978: 113, 126–127, 271, 372–374). Ennek megfelelően például a magyar nyelv történetét bemutató munkákban a nem az alapnyelvből magyarázott szavak alakulásmódja kizárólag mint szókészlet-változási kérdés vetődik fel, grammatikai folyamatokról pedig szó sem esik (vö. SZILÁGYI N. 1999: 359–360). (Ezekről részletesen lásd FEHÉR 2011a: 111–114).

²⁹ Ennek a metodológiai alapelvnek a nyelvtudományban való jelenlétéhez a magyar szakirodalomból vö. például É. KISS munkáját (1998: 25–28). A chomskyánus nyelvészet analógiapárti kritikáját, valamint az ebből kibontakozó vitát pedig — ugyancsak hazai vonatkozásban — KÁLMÁN 1996, majd BRÓDY 1996.

³⁰ Az analógia mint a nyelvi folyamatokat magyarázó elv leginkább a 19. század végén kialakult újgrammatikus történeti nyelvészetben volt és van jelen. A saussure-i strukturalizmus 20. század eleji megjelenésével az analógia nyelvelírási jelentősége összességében csökkent, és jórészt a szabályok szolgálatába állítódott, így a „rendszerkényszer” szinonimájaként még manapság is főként a rendhagyó nyelvi elemeknek a szabályosakhoz való visszatérését jelenti. (A kérdéstről kifejtett formában lásd FEHÉR 2013b, innen is különösen 64–68.)

nyelvelsajátításban betöltött szerepük szempontjából hangsúlyosnak az előbbit tartja. Ez egy egészen sajátos nyelvelsajátítási forгатókönyvhöz, az Elvek és Paraméterek (Principles and Parameters) koncepcióhoz vezet, amelyben az „elvek” a minden nyelvre érvényes általános struktúrákra utalnak, a „paraméterek” pedig ezek nyelvsajátos megjelenési lehetőségeire vonatkoznak.³¹

CHOMSKY szerint a világ nyelveinek általános, absztrakt grammatikája, az ún. Univerzális Nyelvtan (Universal Grammar) egy innát, zárt és autonóm logikai-mondattani meghatározottságú modulként értelmezendő. Mivel a nyelvelsajátítás fokozatosságának tényét nehéz volna vitatni, ennek a grammatikának az elvei az elképzelés szerint szakaszosan bontakoznak ki. Vannak köztük olyanok, amelyek a kezdetektől fogva aktívak, de vannak olyanok is, amelyek a születéstől mintegy kétéves korig „alszanak” az agyban, majd hirtelen működésbe lépnek. Ez utóbbi egyfajta érés, ami lényegét tekintve éppen olyan, mint a többi biológiai maturáció: genetikailag programozott, tehát teljes mértékben független a környezeti hatásoktól (vö. CHOMSKY 1959, 1965/1986: 135–140, 1967/1974 stb.).³²

A grammatika e tulajdonságainak két alapvető elméleti következménye van. Egyrészt az Univerzális Nyelvtan (egy-egy komponenseinek) aktiválódásakor a rejtett nyelvtani kompetenciából egy csapásra nyilvánvaló grammatikai tudás lesz, amit ezután a nyelvelsajátító egyén tanulással — értsd: az adott nyelvre érvényes paraméterek kiválasztásával — anyanyelvére vonatkoztat, azaz megfigyeli, hogy a környezetében használt nyelvben elöl vagy hátul állnak-e a toldalékok, van-e számbeli egyeztetés a melléknév és főneve között stb. Másrészt az Univerzális Grammatika tétele olyan velünk született szabályrendszer feltételez, amely nem engedi meg a hibás szabályáltalánosításokat, a téves hipotézisek felállítását.³³

Azt, hogy a gyerekek nyelvében mégis előfordulnak agrammatikus elemek, CHOMSKY annak tulajdonítja, hogy a paraméterbeállítások még nem történtek meg. Úgy gondolja, a nyelvelsajátító babák még nincsenek kellő számú nyelvi tapasztalat birtokában, amikor először kísérletet tesznek arra, hogy az addig statisztikai úton megtanultakból kiemeljék az anyanyelvükre alkalmazható absztrakt logikai formulákat: a szabályok szempontjából „kivételes” alakokat a mentális szótárnak kellene tartalmaznia, ami azonban még hiányos (BABARCZY–LUKÁCS–PLÉH 2014: 465). Szerinte tehát valójában ez a lexikális korlát és nem egy téves szabálytúláltalánosítás vezet az olyan formák megjelenéséhez, mint mond-

³¹ Az Elvek és Paraméterek teória magyar (nyelvű) összefoglalását lásd például É. KISS–SZABOLCSI 1992: 27–29, PINKER 1994/2006: 109–110, GYÓRI 2008: 31–32, BABARCZY–LUKÁCS–PLÉH 2014: 461–466, GERVAIN 2014: 141–143.

³² Az Univerzális Nyelvtan gondolatának jelenlétét a hazai formális nyelvészetben kognitív jelenségekre vonatkoztatva lásd például KIEFER vagy É. KISS írásaiban (KIEFER 1992: 7, 13, É. KISS 1998: 23–25, 38–39), ennek a nyelvészeti elképzelésnek az interpretációit, valamint a kritikáját a magyar pszicholingvisztikai és kognitív pszichológiai szakirodalomban pedig LUKÁCS–PLÉH 2003: 532–533, PLÉH–LUKÁCS 2003: 496–500, JANCsó 2004: 129–130, 133–134, MAGYARI 2005: 452.

³³ Ezekhez lásd még BABARCZY–LUKÁCS–PLÉH 2014: 461, 465.

juk az angol *go* ige múlt idejű *went* alakja helyett a **goed* vagy a magyar *hó* főnév *havat* tárgy esete helyett a **hót*. A gyerekek mentális szótárak növekedésével az Univerzális Nyelvtanuk paraméterbeállításait elvégzik, így körülbelül hatéves korukig tart, míg valamennyi nyelvsajátos algebrai struktúra kiválasztásához el nem jutnak (CHOMSKY 1981, vö. még 1983: 1–16).

Akár számolunk téves szabályáltalánosításokkal, akár nem, a kognitív fejlődés menetét a teljesítmény és az idő vonatkozásában ábrázolva jól látszik, hogy ebből a nyelvsajátítási foratókönyvből nem egy, hanem két, eltérő jelentőségű és formájú függvény rajzolódik ki. Abban, hogy itt éppen egy ilyen kétutas modellt kapunk, nincsen semmi szokatlan: CHOMSKY eleve egy dichotómiából, a nyelvtan–szótár kettősségből indult ki, a szembeállító pár tagjainak pedig mind fontosságuk, mind alakulásmódjuk tekintetében más-más tulajdonságokat előfeltelezett. Ennek megfelelően nála a nyelvi folyamatoknak van egy hangsúlyosabb, grammatikai vonala és egy mellékesebb, lexikai ága, az előbbi függvénye lépcsőzetes, az utóbbié lineáris.

A nyelvtan lépcsős függvényének kezdete az Univerzális Nyelvtan először aktiválódó komponensének megjelenéséhez, további fázisai pedig a többi modul működésbe lépéséhez, illetve ennek részeként a paraméterbeállításokhoz köthetők. Konstans egyenesekből áll, és ugrásokkal emelkedik, de csak néhány szakaszon át, mert aztán egy adott magasságban végérvényesen stagnál. Ez utóbbi akkor következik be, amikor az idő telik ugyan, de a nyelvtani teljesítmény már nem változik: a szabálykiválasztó paraméterbeállítások befejeződnek, az adott grammatikai szerkezetek használata pedig rögzül.

A szótanulás lineáris függvénye ehhez képest egészen más, és jóval egyszerűbbnek is tűnik: minél több idő telik el, az adott személy annál több szót ismer. Ez pusztán egy hét-nyolc hónapos kor tájékán induló, időarányosan emelkedő egyenes, amelyen nem találunk semmiféle olyan részt, amit a többihez viszonyítva fordulópontként kiemelhetnénk. Így ez a lexikai bővülési folyamat kognitív jelentőségét tekintve elmarad a grammatikai kompetencia heurisztikája mögött.

25. Bár a chomskyánus elképzelés manapság is eléggé népszerű mind a nyelvészek, mind a kognitív pszichológusok körében,³⁴ annak, hogy nem problémátlan, a legszembetűnőbb jele talán éppen az, hogy a lépcsős és lineáris függvények kétutas modellje egyáltalán *n e m i l l e s z k e d i k* azokhoz a görbékhez, amit a gyereknyelvi tesztek adatai alapján kapunk.³⁵ Mindazonáltal erre tulajdonképpen számítani lehet, ha figyelembe vesszük, hogy CHOMSKY nyelv-

³⁴ A chomskyánus elképzelés legismertebb továbbgondolója, részletes kifejtője PINKER. Közismert nyelvsajátítás-monográfiái (1984, 1994/2006) azóta iskolateremtővé lettek a pszicholingvisztikában.

³⁵ Az anyanyelv-elsajátítás S, illetve U alakú görbéiről a morfoszintaktikai tanulás kapcsán részletesebben lásd FEHÉR 2013a, innen is főként 152–155. Ugyanezekről általánosabban, a grammatika hálózatalvú felfogását bemutatva FEHÉR 2013b, különösen 74–76.

elsajátítási koncepciója háttérében a nyelvtani kompetencia és a lexikális megismerés merev elkülönülésének előfeltevése áll. Ez a dichotómia ugyanis mindkét elemében nyugtalanító kérdéseket vet fel: a tudás és tanulás szétválasztása éppúgy, mint a grammatika és a lexikon elhatárolása.

CHOMSKY a tudást a tanulástól egyfelől az innátizmus tana révén kívánja elválasztani, eszerint a világ nyelveinek összes lehetséges nyelvtani szerkezete logikai formulák formájában genetikailag adott, minden mást pedig szociális úton szerzünk. Noha elviekben az így meghúzott határvonal kristálytisztának látszik, a gyakorlatban korántsem ennyire egyértelmű a helyzet.

Koraszülöttekkel végzett mérésekből, valamint a terhesség alatt anyanyelvi ingerekkel trenírozott újszülöttek kísérleti vizsgálataiból tudjuk, hogy már magzati korban is feltételeznünk kell nyelvsajátos tanulást.³⁶ Ha pedig ez így van, aligha lehet megmondani, mi az, amit a géneknek, és mi az, amit a környezetnek köszönhetünk. Úgy tűnik, a kontextus a fejlődésnek minden fokán szerepet játszik: a genom nem valamiféle program, hanem inkább egy sor megkötés, amelyek bizonyos kereteken belül teret engednek a változatosságnak — már magzatoknál is. A genotípus és a fenotípus közt fennálló rés miatt tehát a születést olyan értelemben határnak tekinteni, ahogy azt CHOMSKY teszi, nemcsak hogy önkényes, de meglehetősen problematikus is (MÜLLER 1996).³⁷

CHOMSKY nyelvsajátítás-koncepciójában a tudás és a tanulás az innátizmus tételének következményeként még egy másik szempontból is elkülönül, ám ez a distinkció sem tűnik az előzőnél egyértelműbbnek. Miután úgy gondolja, az emberi nyelvek általános algebrai formulái velünk születnek, grammatikai tanulásal nemigen számol, hiszen a biológiailag programozott nyelvtani kompetenciának csupán aktiválódni kell. Érvelése azonban ezen a ponton különösen ellentmondásossá válik.

A grammatikai tudásunk eleve adott, kibontakozásához mégis nyelvi tapasztalatok kellene. Ez utóbbiakat viszont lexikonunkat bővítve, vagyis statisztikai úton szerezzük meg. Ha pedig ez így van, akkor egy eléggé zavarba ejtő paradoxonnal találjuk szemben magunkat. Nevezetesen azzal, hogy egy statisztikai tanulóval épülő szótárból válnak ki végül logikai nyelvtani struktúrák, de úgy, hogy mindeközben a megjelenő grammatika szigorúan algebrai formulákon alapul, vagyis a lexikont jellemző valószínűségi mintázatokról teljesen független.

Azt persze lehetne mondani, hogy ez esetben jobb, ha inkább kétféle kognitív nyelvtant teszünk fel: egy statisztikait és egy algebrait. Az előbbi a lexikális tanulás során fokozatosan jelenik meg, az utóbbi pedig ugrásszerűen, egy kritikus

³⁶ A vonatkozó tesztekhez vö. többek között DECASPER–FIFER 1980, DECASPER–SPENCE 1986, DECASPER és mtsai 1994, LECANUET 1998, BYERS-HEINLEN–BURNS–WERKER 2010. A kísérleti eredmények rövid, magyar nyelvű áttekintését lásd CSIBRA 2003: 256–257.

³⁷ A témához vö. még a magyar szakirodalomból a kognitív tudományok oldaláról JANCsó 2004, nyelvészeti részről pedig FEHÉR 2013d.

pont után. Ez abban a tekintetben nem is túl ésszerűtlen elképzelés, hogy nemcsak egy heurisztikus, hanem egy olyan grammatikai tudást is feltételez, ami már a tanulás menete közben, vagyis folyamatosan épül. Ám ha a nyelvtant ennek értelmében részben statisztikainak, részben algebrainak tartjuk, továbbra is marad két nyugtalanító kérdésünk, amelyek ezeknek a grammatikáknak a geneziséét, illetve evolúciós hatékonyságát érintik.

Egyrészt továbbra sem világos, hogy hogyan jöhet létre egy algebrai nyelvtan egy statisztikaiból. Másrészt azt sem lehet tudni, hogy egyáltalán miért szükséges a korábban jelentkező és a nyelvi tapasztalatok függvényében folyamatosan alakuló valószínűségi grammatika mellé még egy másik, algebrai nyelvtan is, ami ráadásul hirtelen tűnik fel, és a formálódása viszonylag hamar lezárul.

A grammatikai tudás és a lexikai tanulás dichotómiájának következtében nem érhető az sem a CHOMSKY-féle nyelvelsajátítás-elméletben, hogy ha kisgyerek korban, a paraméterbeállítások befejeződésével a kiválasztott nyelvtani formulák használata rögzül, hogyan lehetségesek olyan grammatikai változások, amik nem iskoláskor alatti gyerekektől származtathatók. Mintha itt a tanulástól elválasztott tudás abszolútizálna.

Azt feltételezni, hogy a nyelvtudás kognitív története egy velünk született, jól körülhatárolható és univerzális grammatikai struktúra kompetenciájával indul, a lehetséges algebrai formulák közül való válogatás, így egy adott nyelvtan elsajátítása pedig egy olyan ponttal zárul, amikor már minden egyes grammatikai paramétert megfelelően beállítottunk, csak valami külső és rögzített viszonyítási ponthoz képest lehet. Márpedig egy autonóm és invariáns nyelvi korpusz a természetes nyelvi folyamatokkal aligha hozható összhangba.³⁸

CHOMSKY-nál azonban ez a fajta gondolkodásmód eléggé egyértelmű, mivel explicit formában is tetten érhető.³⁹ Az emberektől függetlenített, homogén és

³⁸ Ez a szemléletmód egyedül abban az esetben volna helyénvaló, ha nyelvészként olyan modellt állna szándékunkban elkészíteni, ami nem feltétlenül a tényleges nyelvi történéseket szemlélteti, csak egy korpusz adatait foglalja rendszerbe egy meghatározott, de az emberi nyelv mechanizmusaival nem mindenképpen egyező módon. Ekkor nemcsak hogy megengedhető, de sok esetben célszerű is, hogy a leírás során használt kategorizáció megválasztása ne a nyelv tulajdonságaitól, hanem a konkrét alkalmazott kutatási céltől (például beszéd felismerők, helyesírás-ellenőrzők, fordítóprogramok készítésétől) függjön.

Csak hogy ilyen szándékot a modern nyelvtudomány mintegy száz-százötven éve alatt pusztán a bloomfieldi deskriptív iskola és az újabb, kifejezetten disszeminációs célzatú számítógépes nyelvészet fogalmazott meg. Ezeket leszámítva azonban minden lingvisztikai irányzat a valós nyelvi folyamatok megragadásának igényével lépett fel, és erről mind a mai napig nem is mondott le. (A szándék és a megvalósítás eltérésének sajátos módszertani problémáiról részletesen — a beszédmű és a beszédtevékenység elkülönítése kapcsán — lásd SZILÁGYI N. 2004: 47–132.)

³⁹ Meg kell jegyezni, hogy ezzel CHOMSKY a nyelvészetben koránt sincs egyedül. A nyelv fogalmának abszolútizálása — különböző formában és mértékben — a jelenlegi nyelvtudományi irányzatok szinte mindegyikénél fellelhető. Elég, ha ennek kapcsán csak a 20. századi nyelvészet ikonikus alakjának, SAUSSURE-nek a langue–parole dichotómiájából az első tag jellemzését idézzük: „A nyelv úgy él a társadalomban, mint a valamennyi agyban tárolt lenyomatok összege, körülbelül úgy, mint egy szótár, amelynek azonos példányaikat szétosztották az egének között. Ez te-

pontosan definiálható nyelvtan nála a folyamat elején a szerkezeti innatizmus tanában ölt testet, a végén pedig a szociális kontextusából kiemelt „ideális” beszélő-hallgató⁴⁰ elméjében⁴¹ jelenik meg.

26. Miközben a tudás és a tanulás, valamint a nyelvtan és a szókészlet bármiféle elválasztása problémákat vet fel, nem látszik semmi akadály annak, hogy egy olyan *statisztikai grammatika*val számoljunk, amely fokozatosan épül, és amelynek a szótár éppúgy inherens része, mint a morfológia és a szintaxis.⁴² Azt, hogy ez nem is annyira képtelen elgondolás, furcsa módon többek között a kognitív pszichológiai kísérletekben használt stimulusok nyelvi kategorizálásának kaotikussága jelzi.

SAFFRANÉK (1996), MARCUSÉK (1999) és PEÑÁÉK (2002) méréseikhez egyaránt olyan nyelvi egységeket konstruáltak, amelyek három, mássalhangzó + magánhangzó szerkezetű szótagból álltak (így például *golabu*, *bidaku*, *padoti*; illetve *gatiga*, *linali*, *nigini*, valamint *puliki*, *puraki*, *pufoki*). Csakhogy SAFFRANÉK ezekkel a szótanulást, MARCUSÉK és PEÑÁÉK viszont a nyelvtani szerkezetek elsajátítását kívánták vizsgálni. Mindez ráadásul úgy történt, hogy egyedül MARCUS és kollégái tekintették a műnyelvük egységeit mondatoknak, hiszen PEÑA és társai SAFFRAN csoportjához hasonlóan szavakként definiálták az általuk használt szótaghármassokat.

Persze egy, a szótárt is magában foglaló statisztikai grammatika ötlete csak abban az esetben lehet releváns elképzelés, ha az, hogy a lexémák felépítése sta-

hát olyasvalami, amely mindegyikükben külön-külön van meg, mégis mindannyiukban közös, és a letéteményesek akaratán kívül áll” (1916/1997: 49).

⁴⁰ „A nyelvelmélet [értsd: a chomskyánus nyelvészet] elsődlegesen egy teljesen homogén beszélőközösség ideális beszélőjével-hallgatójával foglalkozik, aki tökéletesen tudja nyelvét, és akit nem befolyásolnak olyan grammatikai szempontból érdektelen körülmények, mint az emlékezet korlátai, kihagyásai, a figyelem és az érdeklődés elterelődése és a hibák (legyenek azok véletlenszerűek vagy jellegzetesek), amikor tényleges performanciájában alkalmazza tudását” (CHOMSKY 1965/1986: 115, vö. még 1986: 19–24).

⁴¹ „Elképzelhető egy olyan ideális, homogén nyelvközösség, melyben nincsenek stiláris vagy nyelvjárási változatok. Azt is feltételezhetjük, hogy a nyelvtudás egyöntetű módon reprezentálódik a beszélőközösség minden egyes tagjának az elméjében” (CHOMSKY 1980: 219–220, magyarul idézi BEZECZKY 2002: 47).

⁴² A felvetés lényegi vonását, a szókészlet nyelvtaniságát érzékeltethetnénk akár a „lexikongrammatika” elnevezéssel is (így tettem például magam is; FEHÉR 2008b: 44–45, 2013a: 157). Félő azonban, hogy ez éppen annyira lehetne félrevezető, mint amennyire szemléletes. Hiszen ugyanekkor alapunk volna arra is, hogy ezt a jelenséget a fordított „grammatikalexikon” címkével emlegessük, sőt arra is, hogy ezek közül egyiket se használjuk. Mivel itt arról van szó, hogy a szókészlet és nyelvtani jellemzők bármiféle megkülönböztetése aggályos, e jelenség kapcsán könnyen megtévesztővé válhat minden olyan jelentéstömörítő vagy minőségjelzős kifejezés, amely erre az azonosságra nem tud kellő egyértelműséggel utalni.

Ezért ezen lehetséges szakszavak közül terminusnak ezúttal inkább egyiket sem választjuk. Már csak azért sem, mert erre nem is igen lesz szükségünk. Ahogy az a későbbiekben látszani fog, a grammatika fogalma még ennél is tágabb: az itt felvázolt nyelvtani modellbe további, szokásosan nem grammatikai jegyként emlegetett tulajdonságok is mint szerves rendszerösszetevők kerülnek be.

tisztikailag strukturált, nemcsak a tesztek mesterséges ingeranyagára, de a természetes nyelvek szókészletére is áll. Márpedig a jelek szerint ez a valódi szavakra is vonatkozik: ezek elemei sem random módon kapcsolódnak egymáshoz, vagyis kombinációik meghatározott valószínűségi mintázatokkal írható le.

Olyannyira így van ez, hogy SAFFRAN és munkatársai (1996) a természetes nyelveknek kifejezetten ebből a strukturális jellemzőjéből indultak ki, amikor megtervezték műnyelvi kísérletüket. Az algebrai–statisztikai tanulás vita szempontjából kissé ironikus, hogy ezt éppen CHOMSKY tanárának, HARRISnek az egyik tanulmányára (1955) hivatkozva tették. SAFFRANék amellet érveltek, hogy két szomszédos szó esetében a szó belseji együttjárás gyakorisági mutatók magasabbak, mint azok, amik az első szó végi és a második szó eleji elemek között vannak, a szavak közti szünetek híján tehát ezekre támaszkodva tudjuk megállapítani egy hangsorfolyamban a s z ó h a t á r o k a t .

Példájjukkal élve: az angol *pretty#baby* szekvencia⁴³ esetében az átmeneti valószínűség a *pre*-ből a *tty*-be vagy a *ba*-ból a *by*-be nagyobb, mint a *tty*-ből a *ba*-be, mivel mind a *pretty*, mind a *baby* többször fordul elő a *prettybaby* hangsornál. Ez egész egyszerűen azért van így, mert az előbbieket nemcsak egymás szomszédságában, hanem önmagukban (*pretty*, *baby*), de legfőképpen más szavak társaságában is megjelennek (*pretty#girl*, *cute#baby* stb.) (vö. SAFFRAN–ASLIN–NEWPORT 1996: 1927).⁴⁴

Annak, hogy a szóhatárokat együttjárás gyakorisági mutatók alapján ismerjük fel, van egy igen előnyös tulajdonsága: olyan rugalmas rendszert hoz létre, amely CHOMSKY Univerzális Nyelvtanával szemben a nyelvi tapasztalatoknak megfelelően egyrészt változatos lehet, másrészt pedig változni tud. Miközben a nyelvtudományi munkák rendre problémába ütköznek a szó fogalmának meghatározását illetően,⁴⁵ a statisztikai alapú megközelítés meglepően egyszerű választ ad a kérdésre.

A nyelvészeti szakirodalomban az alapvető nehézséget régóta az összetett szavak ügye jelenti. A gondot az okozza, hogy a szóösszetételek (*képeskönyv*) és a szószerkezetek (*képes#könyv*) elkülönítésére mindenképpen egy egyértelmű és abszolút módon érvényesíthető elvi kritériumot kívánnak találni,⁴⁶ ám a gyakor-

⁴³ A # jel itt és a továbbiakban a szóhatárt jelöli.

⁴⁴ Lásd továbbá BABARCZY–GÁBOR–BÁLINT 2014: 414–417, LUKÁCS 2014: 524–529.

⁴⁵ A próbálkozásokról és ezek paradoxonjairól magyar vonatkozásban és nyelvészeti oldalról néhány éve részletes elemző áttekintést adtam (FEHÉR 2007, 2008a). Ennek a tanulmánynak az utolsó részében egyébként egy az ittenivel teljesen összeegyeztethető javaslatot is körvonalaztam, de azt akkor még nem bontottam ki részleteiben, és nem kapcsoltam közvetlenül a statisztikai grammatika(tanulás) témaköréhez sem (2008a: 60–67).

⁴⁶ Ez az elv a legnagyobb definíciós kihívást jelentő esetekben, vagyis a minőségjelzős alakulatoknál főként a szemantikai transzparenciához kötődik. „A lexikai szó a meghatározás szerint olyan nyelvi elem, amelyben az egész jelentése nem vezethető le, nem rakható össze a részeinek a jelentéséből, pl. [...] a *képeskönyv* összetétel a *képes könyv* szószerkezettől [...] különbözik, hiszen az előbbi egy bizonyos könyvfajtát jelöl, míg az utóbbi a *kép nélküli könyv* kifejezéssel áll szemben” (KENESZI 2000: 78–82).

lat azt mutatja, hogy a konkrét esetek megítélése relatív és többértelmű, hiszen az emberek a szóhatárokat gyakran egymástól eltérően húzzák meg (például a *házinyúl* hangsornál egyesek *házinyúl*-ként, mások *házi#nyúl*-ként). Ez a probléma ugyanakkor a statisztikai szófogalom kapcsán fel sem merül. Ezt az értelmezést követve a „mi a szó?” helyett a „mi egy szó?” kérdésre keressük a választ, így természetesnek vesszük, hogy az emberek különféle nyelvi tapasztalatokkal rendelkeznek, ezért korpuszaikban a hangsorok együttjárási gyakorisági mutatói nem lehetnek tökéletesen egyformák.

Azt, hogy változatosság és változás valóban azokon a pontokon várható, ahol az átmenet-valószínűségek is heterogének és dinamikusak, jól mutatja többek között SOSA és MACFARLANE egyik gyakran idézett pszicholingvisztikai tesztje (2002). A kísérletvezetők angol anyanyelvű felnőttekkel egy szómonitorozási feladatot végeztek el, amelyben az *of* felismerésének reakcióidejét mérték különböző kollokációkban, így magas (*kind of, lot of, one of* stb.), közepesen magas (*couple of, part of, most of* stb.), közepesen alacsony (*care of, because of, kinds of* stb.) és alacsony gyakoriságú (*sense of, piece of, sums of* stb.) előfordulások esetén. Arra kérték a résztvevőket, hogy amikor az *of*-ot hallják, domináns kezüket használva azonnal nyomjanak meg egy adott billentyűt. Kiderült, hogy a válaszadás gyorsasága éppen az eloszlási mutatókat követi: úgy tűnik, minél gyakoribb egy kollokáció, annál nehezebb annak pusztán egyetlen elemét detektálni, mert annál valószínűbb, hogy a személyek az egész hangsort egyetlen szóként kezelik.⁴⁷

27. A szóhatárok átmenet-valószínűségi alapú kijelölése azzal jár, hogy a nyelvtudományban jellegzetesen nem nyelvtani tényezőként számon tartott *f o n o t a x i s t* is grammatikai természetűnek kell minősítenünk. E felfogásban a fonotaktikai struktúrák minőségileg nem, csak statisztikai bonyolultságukat tekintve különböznek a morfoszintaktikai mintázatoktól. Tehát a fonotaxis, a morfológia és a szintaxis voltaképpen egyazon nyelvtani jelenség eltérő komplexitású részei, amelyek egy folytonos skálán helyezhetők el, és nincs köztük éles határ.

Míg a klasszikus szabálynyelvtanokban a fonotaxis egyáltalán nem kap helyet, egy statisztikai modellben nemcsak hogy szerepe van, de egyenesen a grammatika egyik alappilléreként funkcionál. Nem véletlen, hogy a gyereknyelvi kutatások tanúsága szerint a csecsemők különösen érzékenyek a hangsorok szótagszerkezetére.⁴⁸ Tulajdonképpen már a szavak elsajátítása is csak úgy magyarázható, ha a fonotaxissal mint statisztikai grammatikai tényezővel számo-

⁴⁷ Hasonló mérésekről lásd például SOSA–MACFARLANE szakirodalmi áttekintését (2002: 227–230), a kérdéskör magyar nyelvű rövid összefoglalásához pedig vö. PLÉH–LUKÁCS–KAS 2008: 803–814.

⁴⁸ A szótagmintázatok nyelvelsajátításbeli funkciójáról részletesen és szakirodalmi áttekintéssel lásd FEHÉR 2008b: 28–41. Vö. még MACWHINNEY 2003: 507, 512–513.

lunk, még hozzá egy a szokásosnál tágabb értelemben. Ahhoz ugyanis, hogy a szóhangalakokat, így az anyanyelvünk hangsorszerkezeti mintázatait megismerjük, a szótagok közötti átmenet-valószínűségeket kell megfigyelnünk, ami viszont akkor lehetséges, ha egyáltalán tudjuk, mely szótag-előfordulásokat kell egy szótagtípusnak tekintenünk, és melyeket nem.

A fonetikai-fonológiai szakirodalomban korántsem újdonság az a megállapítás, hogy nyelvünkben vannak olyan akusztikai különbségek, amelyeknek jelentőséget tulajdonítunk (ilyen például a magyarban a *láz* és a *ráz* egy szótagos szavak elején lévő tremuláns–approximáns eltérés), és vannak olyanok, amelyeknek nem (ilyen mondjuk a magyar *inda* és *inğa* szavak⁴⁹ első szótagjait záró nazálisok alveoláris–veláris eltérése). A jelenség első ránézésre eléggé esetlegesnek látszik, hiszen az elkülönítés nem függ a különbség mértékétől, és nem is univerzális.

Az *inda* és az *inğa* szavak első szótagjai épp annyira másak, mint például az *ide* és az *ige* második szótagjai. A *de* és a *ge* minden jellemzője megegyezik, kivéve azt az egyet, hogy az előbbi egy alveoláris, az utóbbi viszont egy veláris zárhanggal kezdődik. Az *ide* és *ige* utolsó szótagjait mégis más-más hangsorokként kezeljük, az *inda* és *inğa* első szótagjait viszont nem. Abban, hogy mely különbségek számítanak mérvadónak, ráadásul nyelvenként, sőt nyelvjárásonként is vannak eltérések. Az angol *sin* 'bűn' és (*to*) *siŋ* 'énekelni' szavak egy-egy szótagja ugyanazzal az alveoláris, illetve veláris zárhanggal végződik, mint az *inda* és *inğa* első szótagjai. Míg azonban a magyarok ez utóbbiakat egyfélének tekintik, az angol anyanyelvűek határozottabb különbséget tesznek közöttük. Mindkét nyelvben muszáj ugyanakkor más-más szótagokként kezelni az *r*-t és az *l*-et tartalmazó szótagokat, egyébként aligha lehetne tudni, hogy *lake* 'tó' vagy *rake* 'gereblye', illetve *láz* vagy *ráz* hangzott-e el. Ezzel szemben egy japán anyanyelvű ezeket nem különítené el, neki nem számítana, hogy a *papulika* 'paprika' vagy a *papurika* 'paprika' hangsort mondták-e. A magyar anyanyelvűek egy jelentős hányada legfeljebb „furecsának” tartja, ha valakitől azt hallja, hogy *égyet-tem*. Mindeközben vannak a nyelvterületnek olyan részei, ahol a nyílt *e*-t és a középzárt *ĕ*-t tartalmazó szótagokat majdhogynem annyira másnak tekintik az emberek, mint egy *a*-s és egy *o*-s szótagot. Ezt mutatja például az is, hogy számukra a *mĕntek* 'ők mentek' és a *mĕntĕk* 'ti mentek' jelentései is épp úgy különböznek egymástól, mint mondjuk a *masni* és a *mosni* szavaké a magyarul beszélők körében mindenütt.

Olyan, mintha mindez egy megoldhatatlannak látszó helyzetet teremtene a szótanulásra nézve. Úgy tűnik ugyanis, hogy a kisgyerekek szavakat csak azután sajátíthatnak el, miután rájöttek arra, milyen hangsoreltérésekre kell érzékeny-

⁴⁹ Mivel az érvelésben jelentősége van, a veláris nazális jelölésére itt és a továbbiakban a hangtani szakirodalomból ismert *ŋ* jelet használom, egy későbbi példa kapcsán pedig a magyar egyezményes átírás *ĕ*-jét alkalmazom a középzárt palatális illabiális lejegyzésére. A fonetikus átírást minden más esetben viszont az egyszerűség kedvéért mellőzöm.

nek lenniük. Hiszen egyébként honnan tudnák, hogy két adott szótagpéldányt éppen egy vagy két szótagtípusnak kell-e megfeleltetniük. Ahhoz viszont, hogy kiderítsék, milyen különbségek lényegesek, ismerniük kell azokat a szóalakokat, amelyekben ezek előfordulnak. Különben megállapíthatatlan, hogy mely fizikai eltéréseknek lehet egyáltalán szemantikai jelentőségük. Ez a probléma azért is zavarba ejtő, mert tudjuk, hogy ez az egészen fiatal korosztályt, vagyis a csecsemőket kell, hogy érintse: a gyerekek egyévesen első szavaikat ki is mondják, ezt megelőzően pedig már jóval több szót megértének.

Ez a paradoxon ugyanakkor fel sem merül, ha ez esetben is statisztikai tanulást feltételezünk. A beszédkódnak egyetlen olyan tulajdonsága van, ami a szavak ismerete nélkül segíthet a nyelvelsajátító gyerekeknek a releváns és irreleváns akusztikai különbségek szétválogatásában. Ez pedig nem más, mint az egymással kisebb-nagyobb mértékű hasonlóságot mutató szótagok *gyakorisági eloszlása*: ez egy olyan statisztikai mutató, amely azt jelzi, hogy egy adott korpusz elemei gyakoriságuk szerint hogyan oszlanak el a korpusz egészén belül.

Ez az, ami az átmenet-valószínűségekkal együtt árulkodhat a szótanulás előtt álló gyerekeknek arról, hogy például a magyarban az *l*-es hangsorokat ajánlatos elkülöníteni az *r*-es megfelelőiktől, ám az alveoláris *n*-esek és veláris *ŋ*-esek esetében ezt nem kell megtenni. *R*-ek és *l*-ek szinte minden hangkörnyezetben megjelenhetnek, még hozzá a veláris *ŋ*-nél jóval gyakrabban. Az utóbbi viszont kizárólag a *k* és *g* hangok előtt található meg, de ott mindig (így az *inŋa* mellett például a *inŋka* vagy az *anŋol* szavakban) — ezekből az esetekből azonban nincsen túl sok. Mindez azzal jár, hogy olyan szópárok előfordulhatnak, amelyeket csak az *l*-ek és *r*-ek különböztetnek meg (vagyis eltérésükön szavak jelentése múlik), de ugyanez a *ŋ*-ek kapcsán már nem lehetséges. *K* és *g* előtt csupán a *ŋ* változat tűnhet fel, valamennyi egyéb hangkörnyezetben pedig alveoláris *n*-et lehet hallani.

A gyakorisági eloszlásnak a nyelvelsajátításban betöltött szerepét szemlélteti KUHL és munkatársai (2006) klasszikusnak számító kísérlete is. A kutatócsoport azt tesztelte, hogy vajon az amerikai, illetve japán csecsemők mikor kezdik elsajátítani azt, hogy hogyan kell viszonyulniuk az *r*-et és az *l*-et tartalmazó szótagokhoz. A kutatók *la*-kat és *ra*-kat játszottak le hat-nyolc, valamint tíz-tizenkét hónapos egynyelvű angol, illetve japán környezetben nevelkedő seattle-i és tokiói csecsemőknek, mégpedig úgy, hogy a csoport egyik fele hosszú ideig csupa *la*-kat hallott, mire egyszer feltűnt egy *ra*, a másik fele pedig éppen fordítva. A fejfordításos paradigmát használva azt vizsgálták, hogy az új inger megjelenésekor a csecsemők az oldalirányba elhelyezett hangforrás és játék felé fordítják-e a tekintetüket, vagyis felfigyelnek-e a változásra, így a két szótagot elkülönítik-e egymástól.

A teszt során a bemutatott hanganyagra a hat-nyolc hónapos amerikai és japán babák nagyjából egyformán reagáltak, és a *la-ra* szótagok között hatvanöt százalékban tettek különbséget. A tíz-tizenkét hónaposak eredményei azonban

már jelentősen elváltak egymástól. Miközben a megkülönböztetésben az amerikai csecsemők teljesítménye ekkorra hetvennégy százalékra emelkedett, a japánoké hatvanra esett vissza. Mindkét csapatnál épp úgy indult meg a változás, ahogy ez az anyanyelvük szavainak elsajátítása szempontjából a lepraktikusabb volt, azaz ahogyan ez a környezetük nyelvének gyakorisági eloszlási mutatóiból következett.

Az angol *r*-es és *l*-es szótagokhoz hasonlóak a japánban is előfordulnak, csak más arányban. Az angolban nemcsak hozzávetőlegesen kétszer gyakoribbak, mint a japánban, de akusztikailag is eltérő szerkezetűek. Ha készítünk egy függvényt, amelyen elhelyezzük a két nyelv összes *r*-et és *l*-et tartalmazó szótagját úgy, hogy az *x* tengely két végpontján az egymástól lateralitás tekintetében leginkább különböző szótagok legyenek, az *y* tengely pedig a gyakoriságot jelölje, akkor az angol példák a széleken rajzolnak ki egy-egy haranggörbét, míg a japánok egyet középben. Mindez pedig azt mutatja, hogy ezek között a japán és angol szótagok között akusztikailag vannak átfedések, de csak a haranggörbék szélein: a legtipikusabb ejtésformát reprezentáló szótagok eloszlása a lateralitás szempontjából igen különböző. A japán szótagok valahol az angol centrális *r*-et és laterális *l*-et tartalmazó szótagok között helyezkednek el, és nagyjából fele annyi van belőlük, mint az angol *r*-es és *l*-es szótagokból együttvéve. Az pedig, hogy centrálisak vagy laterálisak-e egy-egy szóban, nem meghatározható: néha ilyenek, néha olyanok, de legtöbbször valahol félúton vannak a centrális és a laterális képzés között.

28. Az a tény, hogy a beszélt nyelvben van koartikuláció, és számolhatunk egyéni hangszínnel is, a megismerés statisztikai menetét már önmagában determinálja. Ezek következtében az emberi nyelvben két szótag fonetikája sohasem lehet tökéletesen ugyanolyan, az egy típusba tartozó példányok valójában csak hasonlítanak egymásra, de semmiképpen nem egyformák. Mivel az algebrai szabályok egyértelmű behelyettesítéseken alapulnak, a szótagpéldányok típusokba rendezésekor használhatatlannak bizonyulnak. A statisztikai analógia viszont ebben az esetben is jól működtethető mechanizmus, mivel nem más, mint egy valószínűségi mintázat, ami egy bizonyos mértékű rokonságot, nem pedig teljes azonosságot vár el.

Mindez ráadásul a nyelvelsajátítás tágabb forgatókönyvéhez is illeszkedik, és a kognitív nyelvi rendszer egészét egy olyan *e g y s é g e s k e r e t b e n* képes kezelni, amiben a fonotaxison túl grammatikai jelenségként értelmeződhet minden, amit hagyományosan fonetikai jegyként, de nem nyelvtani jellemzőként szoktunk számon tartani. A ritmus, a hangsúly és dallam semmiképpen nem algebrai formulák, statisztikai értelemben viszont tekinthetők olyan valószínűségi kombinatorikus mintázatoknak, amelyek az anyanyelv jellegzetes (szupraszegmentális) grammatikai szerkezeteit testesítik meg. Márpedig nagyon úgy tűnik, az anyaméhben éppen ezek elsajátítása kezdődik el, méghozzá statisztikai tanulással, vagyis a külső nyelvi tapasztalatokra alapozva.

Az anya teste a magzatvízzel egy speciális szűrő, ami nem egyformán engedi át a hangsorok minden akusztikai tulajdonságát: az alacsonyabb frekvenciákból többet, a magasabbakból kevesebbet. Ez a nyelvi ingerek szempontjából azzal jár, hogy az ily módon szűrt beszéd prozódiai és intonációs mintái markánsabban, míg fonetikai aspektusai csak kevésbé érzékelhetők a magzat számára.⁵⁰ Ez pedig teljesen egybevág azzal az általános mérési eredménnyel, hogy az újszülöttek anyanyelvüket képesek elkülöníteni más idegen nyelvektől, de csak azoktól, amelyek szupraszegmentális jegyeikben különböznek egymástól.⁵¹ És megfelel azon kísérletek tanúságának is, amelyek arra utalnak, hogy ha a kismama többnyelvű közegben él, a születendő csecsemőnek már magzati korában meg-alapozódhat a kétnyelvűsége.⁵²

29. Bár minden eddigi érv a statisztikai grammatika mellett szól, ellenvetésként mondhatjuk, hogy a kognitív pszichológiai és a nyelvtudományi szakirodalomban vannak olyan jelek, amelyek mégsem egy ilyen nyelvtanra mutatnak, hanem egy a chomskyánus koncepcióval összeegyeztethető kettős modellre. Úgy tűnik, hogy a felnőttekkel végzett pszicholingvisztikai nyelvfeldolgozási kísérletek eredményei egy kétutas, *s z a b á l y o s – k i v é t e l e s*, illetve *a n a l i t i k u s – h o l i s z t i k u s* kognitív rendszerre utalnak. Ez még ráadásul a nyelvtudományi tradícióval is harmonizál, hiszen a közösségi nyelv morfoszintaxisát a nyelvészeti munkák is a *s z a b á l y o s – a n a l ó g i á s*, valamint a *s z a b á l y o s – r e n d h a g y ó* párokkal írják le. Ha azonban ezeknek a méréseknek a módszertani körülményeit is alaposabban szemügyre vesszük, könnyen kiderülhet, hogy igazából ezek sem cáfolják egy statisztikai grammatika létjogosultságát.

A pszicholingvisztikai tesztek tervező szakemberek általában CHOMSKY felfogása szerint haladnak, hiszen egyrészt a szabályok létezését mintha előfeltételeznék, másrészt pedig mintha eleve kizárnák azt, hogy lehetséges másfajta produktív nyelvtani művelet.⁵³ Ezt már a nyelvi rendszer kétféleségének jellegére utaló szakszavak is szemléltetik: ha a szabályok mellett más termékeny nyelvtani művelettel is számolnának, akkor még egy kétutas modellben sem a „kivételes” és a „holisztikus” jelző volna szembeállítva a „szabályossal”. A kísérletvezetők ennek megfelelően rendszerint a nyelvi példáikat is úgy csoportosítják, hogy már előzetesen reguláris (*tanul, tanult* stb.) és kivételes (*megy, ment* stb.) típusokat különítenek el, majd pedig a mérési eredményeket is kategóriánként értékelik.⁵⁴ Ez az eljárás — függetlenül a valós nyelvi folyamatoktól — mintegy

⁵⁰ LECANUET 1998. Vö. még CSIBRA 2003: 256–257.

⁵¹ Ehhez lásd MEHLER és mtsai sokat idézett kísérletét (1988).

⁵² BYERS-HEINLEN–BURNS–WERKER 2010.

⁵³ A kétutas modellekről e kérdéskör kapcsán magyarul lásd PLÉH 2000: 957–966, 969, 999–1001, 2006b: 736–737, 739, JUHÁSZ–PLÉH 2001: 13–14, 19–32, LUKÁCS 2001, 2008: 122–135, THUMA–PLÉH 2001, PLÉH–LUKÁCS–KAS 2008: 824–827.

⁵⁴ A morfoszintaktikai kísérletek összefoglaló áttekintéseit magyar nyelven lásd például LUKÁCS 2001, PLÉH 2006b: 739–741, PLÉH–LUKÁCS–KAS 2008: 826–827.

előre kijelöli azt az értelmezési keretet, miszerint van eltérés a szabályosnak és a rendhagyónak tartott formák között.⁵⁵

Hasonló probléma vethető fel az adatok gyakorisági értékeinek megállapítása kapcsán is. Noha a gyakorisági hatások a tesztekben nagyon robusztusnak látszanak⁵⁶, ott, ahol nem található összefüggés az eloszlási és a grammatikai tényezők között, többnyire két dolog figyelhető meg. Egyrészt az, hogy a kvantitatív mutatókat a beszélt nyelvtől távol álló, írott korpuszokból számolják. Másrészt pedig az, hogy az emberek tényleges nyelvi tapasztalataihoz igazodó relatív token-gyakoriság helyett az abszolút típusgyakoriságot veszik alapul anélkül, hogy ezt az értékeléskor figyelembe vennék. A kétféle gyakoriság közti különbség korántsem lényegtelen: egy elem előfordulásainak (mondjuk a *tesz* hangsor megjelenéseinek) a száma az egyén beszélt nyelvében lehet magas annak ellenére, hogy a grammatikai csoportnak, amelyhez az adott elem tartozik, csak kevésfajta tagja van (például a *tesz* esetében az *sz*-et *v*-vel váltakoztató igetőtípusnak mindössze hét).⁵⁷ Ahogy a PLÉH–LUKÁCS–KAS szerzőhármas is megjegyzi: „A beszélt nyelv gyakorisági eloszlásai azonban minden bizonnyal eltérnek az írott nyelv gyakorisági eloszlásaitól, és a beszélőt ezek inkább befolyásolhatják. A gyakoriságok idővel és az egyéni környezettől függően változhatnak, így a régi korpuszok alapján számolt mutatók nem biztos, hogy jól tükrözik a vizsgálat időpontjára és az egyénre jellemző eloszlásokat” (2008: 810).

Ami pedig a szabályalapú nyelvészeti grammatikaelméleteket illeti, ezek nem a nyelvi folyamatokat tükrözik, inkább az adatokkal mutathatnak egyezést. Analógias jelenségként vagy taxonomikusan listázva pontosan olyan arányban tűnnek fel bennük a „kivételek”, mint amilyen mértékű a korpusz elemeinek a szabályoktól való elhajlása. Az, hogy ezek a modellek egy bizonyos szinten mégis közelítenek az adatok közti tényleges összefüggésekhez,⁵⁸ abból adódik, hogy a nyelvi relativitásnak vannak kisebb mértékben megnyilvánuló (ezáltal könnyen tökéletesen szabályosnak látszó) esetei. Ilyen lehet a magyar morfológiában a palatoveláris illeszkedés, amely ingadozást (*férfinak* ~ *férfinék*, *farmerban* ~ *farmerben*, *konkrétan* ~ *konkréten*) és kategóriaszórást (*hídja*, *íja*, *de híre*, *íve*) csak az elöl képzett, illabiális tövégi vokálisok esetében mutat.⁵⁹ Ennél nagyobb vál-

⁵⁵ Vö. mindehhez MACWHINNEY 2003: 517, PLÉH 2006a: 764, 2006b: 739–740, LUKÁCS 2008: 127–131, 135.

⁵⁶ Ezekről a kísérletekről jó összefoglalást ad magyar nyelven LUKÁCS 2008: 131–134.

⁵⁷ Erről az igetőtípusról a nyelvészeti szakirodalomban részletesen lásd például VELCSOVNÉ 1968: 99, KESZLER 2000c: 97, a kétféle gyakoriság összefüggéseihez pedig kognitív pszichológiai oldalról vö. LUKÁCS 2008: 132.

⁵⁸ Lásd még FISER 2005: 27, 35.

⁵⁹ Klasszikus fonetikai leírásához vö. többek között PAPP 1966: 124–127, KASSAI 1998: 141–142, GÖSY 2004: 251–252, rendszertani megközelítéséhez pedig DEME 1961: 96–98, VELCSOVNÉ 1968: 96–113, NÁDASDY–SIPTÁR 1994: 94–152, SIPTÁR 1999: 302–312, REBRUS 2000: 786–803 stb., vö. ugyanakkor SZILÁGYI N. sok tekintetben újszerű, ún. beszédmű-fonológiai megközelítését (2004: 206–217).

tozatosság és mozgás jellemző ugyanakkor például a tőtípusrendszerre, hiszen a szép számú és elemzési keretenként eltérő alkategória mellett is maradnak (analógiásnak vagy lexikainak tekintett) „kivételek”.⁶⁰ Az ún. tövéghangzó–kötőhangzó–előhangzó fonetikai minőségének szabályalapú leírása pedig — a nagyfokú variabilitás miatt — már különösen bonyolultnak látszik.⁶¹

Nem mellékes az sem, hogy a hétköznapi nyelvtudás természetére vonatkozó egyszerű megfigyeléseink is inkább egy analógiás grammatikára, mintsem szigorú logikai nyelvtani műveletekre utalnak: „a reguláris nyelvek szabályainak konzisztens alkalmazásaival létrehozott példák egyáltalán nem triviálisak az ember számára. Annak eldöntése, hogy egy szimbólumsor egy reguláris nyelvet követ-e vagy sem, nem a kisgyermek számára is egy pillantással eldönthető kérdés, hanem vizsgapélda egy [...] egyetemista számára” (FISER 2005: 34–35). Ezzel összhangban van a pszichológus COSMIDES és TOOBY méréseinek eredménye, miszerint általában véve könnyebben oldunk meg logikai jellegű problémákat, ha azokat a mindennapos konkrét tapasztalataink példáján keresztül mutatják be, mint ha absztrakt változók és képletek formájában tárják eléink (1993). MAJOROS pedig pedagógiai tapasztalatait összegezve már egyenesen arra jut, hogy a diákok matematikai tévedései nagyrészt abból adódnak, hogy az e tudományban használatos elvont fogalmakat a maguk számára nem az ennek megfelelő algebrai tanulási mechanizmussal, hanem a jóval rugalmasabb anyanyelvi típusalkotás analógiájára formálják meg (1998, innen is különösen 279–280).

Mindezek tükrében erősen megfogalmazottnak és egyben lényegtelennek látszik az a formális nyelvészeti oldalról — főként a szintaxis vonatkozásában — gyakran hangoztatott érv is, miszerint kétséges, hogy egy nem szabályalapú mechanizmussal el lehet sajátítani egy nyelv mondattanát. A kérdésfelvetés azért *i r r e l e v á n s*, „mert csak nagyjából igaz, hogy az ember szintaktikai szabályokat tud és használ, és ennek a nagyjábóloli használatnak sokkal egyszerűbb megvalósíthatósági formái is vannak, mint egy reguláris nyelv szintaxisának tökéletes elsajátítása. Így aztán nincs szükség arra, hogy kiderítsük, hogyan is »képes« az ember egy reguláris nyelv szintaxisának tökéletes elsajátítására, mert nem is teszi ezt” (FISER 2005: 27). A fő vita-, illetve kutatási kérdés ehelyett inkább az kellene legyen, hogy milyen tényezők miatt és hogyan térünk el preferenciáinkkal a szabályos struktúráktól (FISER 2005: 35).

30. Bízom abban, hogy az algebrai–statisztikai tanulás kognitív pszichológiai vitáját elemezve több, egymással összefüggő klasszikus grammatikamodellézési probléma hátterére sikerült egy új, interdiszciplináris nézőpontból rámutatnom.

⁶⁰ A különféle tótani osztályozásokat lásd például KÁROLY 1961b, BERRÁR 1967: 216–229, KESZLER 2000a: 175–182, 2000b, SÁROSI 2005: 129–138, 352–357, 610–612, 719–720, 800–801.

⁶¹ Az ezzel kapcsolatos jellegzetes érvekhez és álláspontokhoz vö. KÁROLY 1961a: 302–312, BERRÁR 1967: 217–218, VELCSOVNÉ 1968: 92–93, KIEFER 1999: 189–261, 2000: 47–48, KIEFER–LADÁNYI 2000: 156–157, LACZKÓ 2000: 45–46, PRÓSZÉKY 2000: 1045–1046. Ezekről eltérő javaslathoz lásd ugyanakkor SZILÁGYI N. 2004: 218–229.

Azt is remélem, hogy az analógiás nyelvtan mellett felhozott érveim hozzájárulnak ezek mielőbbi és hatékony megoldásához.

Egy statisztikai grammatikaelmélet részleteinek kidolgozása és legfőképpen nagy méretű korpuszokon való kísérleti, illetve számítógépes tesztelése tényleg időszerű volna, mégpedig a két tudományterület együttműködésével. A jelek szerint ugyanis a két diszciplína között nem túl erős a szakmai kommunikáció, miközben mind a modern nyelvtudomány nyelvtanképében, mind a kognitív pszichológia grammatikai modelljeiben az algebrai szabály mint magyarázó elv tulajdonképpen módszeres empirikus ellenőrzés nélkül, előfeltevésként szerepel, a statisztikai analógia meg eleve nincsen számításba véve. Ez utóbbi legfeljebb csak olyan „mostohagyerekként” jelenik meg, amit a strukturális anomáliák, azaz a kivételek kezelése miatt kellett pusztán teoretikus megfontolásokból utólag és a lehető legkorlátozottabb mértékben beiktatni a rendszerbe.

A priori elveken alapuló, egy-egy szűkebb tudományterülethez tartozó kutatások helyett sokkal inkább *interdiszciplináris* és a *posteriori* módszerekkel dolgozó vizsgálatokra volna szükség. Így érhető el, hogy a grammatikai koncepciók kapcsán megszüntessük azt a sajátos helyzetet, amiről volta-képpen már WITTGENSTEIN is ír nyelvfilozófiai munkájában: „Az alapvető tény itt az, hogy szabályokat állapítunk meg, egy technikát rögzítünk egy játék számára, és amikor aztán a szabályokat követjük, nem úgy megy a dolog, ahogy feltételeztük. Hogy tehát mintegy önnön szabályaink ejtenek foglyul” (1952/1992: 82).

Irodalom

- ABLA, DILSHAT–KATAHIRA, KENTARO–OKANOYA, KAUZO 2008. On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience* 20: 952–964.
- ALTMANN, GERRY T. M.–DIENES, ZOLTÁN 1999. Rule learning by seven-month-old infants and neural networks. *Science* 284: 875.
- ASLIN, RICHARD N.–SAFFRAN, JENNY R.–NEWPORT, ELISSA L. 1998. Computation of conditional probability statistics by 8-month-old infants. *American Psychological Society* 9: 321–324.
- BABARCZY ANNA–GÁBOR BÁLINT–SERÉNY ANDRÁS 2014. Nyelvtechnológia és a pszicholingvisztika számítógépes modelljei. In: PLÉH CSABA–LUKÁCS ÁGNES szerk., *Pszicholingvisztika. Magyar pszicholingvisztikai kézikönyv 1–2*. Budapest, Akadémiai Kiadó. 411–441.
- BABARCZY ANNA–LUKÁCS ÁGNES–PLÉH CSABA 2014. A nyelvelsajátítás elméleti modelljei. In: PLÉH CSABA–LUKÁCS ÁGNES szerk., *Pszicholingvisztika. Magyar pszicholingvisztikai kézikönyv 1–2*. Budapest, Akadémiai Kiadó. 445–481.
- BERRÁR JOLÁN 1967. A szóelemek alaktörténete. In: BÁRCZI GÉZA–BENKŐ LORÁND–BERRÁR JOLÁN, *A magyar nyelv története*. Budapest, Tankönyvkiadó. 212–235.
- BEZECZKY GÁBOR 2002. *Metafora, narráció, szociolingvisztika*. Modern filológiai füzetek 58. Budapest, Akadémiai Kiadó.

- BRÓDY MIHÁLY 1996. Van-e tűz Varsóban? Megjegyzések egy recenzióhoz. *BUKSZ* 8/4: 379–381.
- BUIATTI, MARCO–PEÑA, MARCELA–DEHAENE-LAMBERTZ, GHISLAINE 2009. Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *NeuroImage* 44: 509–519.
- BYERS-HEINLEIN, KRISTA–BURNS, TRACEY C.–WERKER, JANET F. 2010. The roots of bilingualism in newborns. *Psychological Science* 21: 343–348.
- CHOMSKY, NOAM 1957/2003. Mondattani szerkezetek. In: CHOMSKY, NOAM, *Mondattani szerkezetek. Nyelv és elme*. Budapest, Osiris Kiadó. 7–132.
- CHOMSKY, NOAM 1959. Review of B. F. Skinner's verbal behavior. *Language* 35: 26–58.
- CHOMSKY, NOAM 1965/1986. A mondattan elméletének aspektusai. In: ANTAL LÁSZLÓ szerk., *Modern nyelvelméleti szöveggyűjtemény 4/1*. Budapest, Tankönyvkiadó. 111–319.
- CHOMSKY, NOAM 1967/1974. Újabb adalékok a velűnkszületett eszmék elméletéhez. In: PAPP MÁRIA szerk., *A nyelv keletkezése*. Budapest, Kossuth Könyvkiadó. 85–96.
- CHOMSKY, NOAM 1975. *Reflections on language*. New York, Pantheon.
- CHOMSKY, NOAM 1980. *Rules and representation*. New York, Columbia University Press.
- CHOMSKY, NOAM 1981. Principles and parameters in syntactic theory. In: HORNSTEIN, NORBERT–LIGHTFOOT, DAVID W. szerk., *Explanations in Linguistics*. London, Longman. 123–146.
- CHOMSKY, NOAM 1983. *Lectures on government and binding. The Pisa lectures*. Studies in generative grammar 9. Dordrecht, Foris Publications.
- CHOMSKY, NOAM 1986. *Knowledge of language. Its nature, origin, and use*. New York, Praeger.
- CHRISTIANSEN, MORTEN H.–CURTIN, SUZANNE L. 1999. Transfer or learning: rule acquisition or statistical learning? *Trends in Cognitive Sciences* 3: 289–290.
- COSMIDES, LEDA–TOBBY, JOHN 1993. Cognitive adaptations for social exchange. In: BARKOW, JOHN–COSMIDES, LEDA–TOBBY, JOHN szerk., *The adapted mind. Evolutionary psychology and the generation of culture*. New York, Oxford University Press. 162–228.
- CRYSTAL, DAVID 2003. *A nyelv enciklopédiája*. Budapest, Osiris Kiadó.
- CUNILLERA, TONI–TORO, JUAN M.–SEBASTIAN-GALLÉS, NURIA–RODRÍGUEZ-FORNELLS, ANTONI 2006. The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Research* 1123: 168–178.
- CSIBRA GERGELY 2003. A kognitív fejlődés idegrendszeri háttere csecsemőkorban. In: PLÉH CSABA–KOVÁCS GYULA–GULYÁS BALÁZS szerk., *Kognitív idegtudomány*. Budapest, Osiris Kiadó. 255–272.
- DECASPER, ANTHONY J.–FIFER, WILLIAM P. 1980. Of human bonding: newborns prefer their mothers' voices. *Science* 208: 1174–1176.
- DECASPER, ANTHONY J.–SPENCE, MELANIE J. 1986. Prenatal maternal speech influences newborns perception of speech sounds. *Infant Behavior and Development* 9: 133–150.
- DECASPER, ANTHONY J.–LECANUET, JEAN-PIERRE–BUSNER, MARIE-CLAIRE–GRANIER-DEFERRE, CAROLYN–MAUGEAIS, ROSELYNE 1994. Fetal reactions to recurrent maternal speech. *Infant Behavior and Development* 17: 159–164.

- DEME LÁSZLÓ 1961. Hangtan. In: TOMPA JÓZSEF szerk., *A mai magyar nyelv rendszere. Leíró nyelvtan 1. Bevezetés. Hangtan. Szótan.* Budapest, Akadémiai Kiadó. 57–119.
- EIMAS, PETER D. 1999. Do infants learn grammar with algebra or statistics? *Science* 284: 435–436.
- ENDRESS, ANSGAR D.–BONATTI, LUCA L. 2007. Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition* 105: 247–299.
- ENDRESS, ANSGAR D.–SCHOLL, BRIAN J.–MEHLER, JACQUES 2005. The role of saliency in the extraction of algebraic rules. *Journal of Experimental Psychology: General* 134: 406–419.
- FEHÉR KRISZTINA 2004. Paradigmák kölcsönhatása az újgrammatikus nyelvkoncepcióban. A magyar történeti személynévkutatás a 20. század elején. *Magyar Nyelvjárások* 42: 5–32.
- FEHÉR KRISZTINA 2007. A szó problémája 1. *Magyar Nyelvjárások* 45: 5–26.
- FEHÉR KRISZTINA 2008a. A szó problémája 2. *Magyar Nyelvjárások* 46: 55–70.
- FEHÉR KRISZTINA 2008b. Grammatika és hangsorminta. *Magyar Nyelvjárások* 46: 21–54.
- FEHÉR KRISZTINA 2011a. A családfamodell és következményei. *Magyar Nyelvjárások* 49: 105–128.
- FEHÉR KRISZTINA 2011b. A nyelv modularizálódó hálógrammatikája és az alkalmazott nyelvtudomány. *Modern Nyelvoktatás* 17/2–3: 87–100.
- FEHÉR KRISZTINA 2013a. A morfoszintaktikai szabályok és a pszicholingvisztika. In: BENŐ ATTILA–FAZAKAS EMESE–KÁDÁR EDIT szerk., „...hogyan legyen a víznek lefolyása...” *Köszöntő kötet Szilágyi N. Sándor tiszteletére.* Kolozsvár, Erdélyi Múzeum-Egyesület. 151–169.
- FEHÉR KRISZTINA 2013b. Analógia és hálózatmodell. In: KÁDÁR EDIT–SZILÁGYI N. SÁNDOR szerk., *Analógia és modern nyelvelírás.* Kolozsvár, Erdélyi Múzeum-Egyesület. 66–87.
- FEHÉR KRISZTINA 2013c. A 20. század eleji újgrammatizmus és a nyelvi változások. *Magyar Nyelvőr* 137: 129–154.
- FEHÉR KRISZTINA 2013d. Innátizmus? — egy pszicholingvisztikai kérdés és a szociolingvisztika elmélete-módszertana. In: KONTRA MIKLÓS–NÉMETH MIKLÓS–SINKOVICS BALÁZS szerk., *Elmélet és empiria a szociolingvisztikában. Válogatás a 17. Élőnyelvi Konferencia — Szeged, 2012. augusztus 30.–szeptember 1. — előadásaiból.* 2013. Budapest, Gondolat Kiadó. 137–162.
- FISER JÓZSEF 2005. Létezik-e külön statisztikai és szabályalapú tanulás az agy vizuális és auditorikus moduljaiban? In: GERVAIN JUDIT–KOVÁCS KRISTÓF–LUKÁCS ÁGNES–RACSMÁNY MIHÁLY szerk., *Az ezerarcú elme. Tanulmányok Pléh Csaba 60. születésnapjára.* Budapest, Akadémiai Kiadó. 26–37.
- FITCH, TECUMSEH W.–HAUSER, MARC D. 2004. Computational constraints on syntactic processing in a nonhuman primate. *Science* 303: 377–380.
- GERVAIN JUDIT 2014. Nyelvek és nyelvtanok. In: PLÉH CSABA–LUKÁCS ÁGNES szerk., *Pszicholingvisztika. Magyar pszicholingvisztikai kézikönyv 1–2.* Budapest, Akadémiai Kiadó. 117–145.
- GERVAIN, JUDIT–MACAGNO, FRANCESCO–COGOI, SILVIA–PEÑA, MARCELA–MEHLER, JACQUES 2008. The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences of the United States of America* 105: 14222–14227.

- GERVAIN, JUDIT–MEHLER, JACQUES–WERKER, JANET F.–NELSON, CHARLES A.–CSIBRA, GERGELY–LLOYD-FOX, SARAH–SHUKLA, MOHINISH–ASLIN, RICHARD N. 2011. Near-infrared spectroscopy. A report from the Donell infant methodology consortium. *Developmental Cognitive Neuroscience 1*: 22–46.
- GOMEZ, REBECCA L.–GERKEN, LOUANN 1999. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition 70*: 109–135.
- GÓSY MÁRIA 2004. *Fonetika, a beszéd tudománya*. Budapest, Osiris Kiadó.
- GYŐRI MIKLÓS 2008. A nyelvi képesség természete. In: CSÉPE VALÉRIA–GYŐRI MIKLÓS–RAGÓ ANETT szerk., *Általános pszichológia 3. Nyelv, tudat, gondolkodás*. Budapest, Osiris Kiadó. 29–58.
- HARRIS, ZELIG S. 1955. From phoneme to morpheme. *Language 31*: 190–222.
- JANCSÓ DANIELLA 2004. A veleszületett egyetemes nyelvtan és a modularitás neurobiológiai megközelítése. Vázlatos áttekintés Ralph-Axel Müller közleménye alapján. In: PLÉH CSABA–GERVAIN JUDIT szerk., *A láthatatlan megismerés*. Budapest, Gondolat Kiadó. 129–142.
- JUHÁSZ LEVENTE–PLÉH CSABA 2001. Többmorfémás szavak megértése a magyarban. In: PLÉH CSABA–LUKÁCS ÁGNES szerk., *A magyar morfológia pszicholingvisztikája*. Budapest, Osiris Kiadó. 11–37.
- KÁLMÁN LÁSZLÓ 1996. Varsóban nem volt tűz. Noam Chomsky: Mondattani szerkezetek; Nyelv és elme. *BUKSZ 8/2*: 130–134.
- KÁLMÁN LÁSZLÓ 2007. Holisztikus szemlélet a nyelvészetben. *Szabad Változók 4*. URL: <http://www.szv.hu/cikkek/holisztikus-szemlelet-a-nyelveszetben>. (2014. 12. 28.).
- KÁLMÁN LÁSZLÓ 2012. Analógiás tanulás asszociatív memóriamoddellel. *Általános Nyelvészeti Tanulmányok 24*: 121–134.
- KÁLMÁN LÁSZLÓ–REBRUS PÉTER–TÖRKENCZY MIKLÓS 2010. Possible and impossible variation in Hungarian. In: KIEFER FERENC–LADÁNYI MÁRIA–SIPTÁR PÉTER szerk., *Current Issues in Morphological Theory. (Ir)regularity, analogy and frequency. Selected papers from the 14th International Morphology Meeting, Budapest, 13–16 May 2010*. Amsterdam, John Benjamins. 23–50.
- KASSAI ILONA 1998. *Fonetika*. Budapest, Nemzeti Tankönyvkiadó.
- KÁROLY SÁNDOR 1961a. A szóelemek általános kérdései. In: TOMPA JÓZSEF szerk., *A mai magyar nyelv rendszere. Leíró nyelvtan 1. Bevezetés. Hangtan. Szótan*. Budapest, Akadémiai Kiadó. 295–312.
- KÁROLY SÁNDOR 1961b. A szótövek. In: TOMPA JÓZSEF szerk., *A mai magyar nyelv rendszere. Leíró nyelvtan 1. Bevezetés. Hangtan. Szótan*. Budapest, Akadémiai Kiadó. 313–331.
- KENESEI ISTVÁN 2000. Szavak, szófajok, toldalékok. In: KIEFER FERENC szerk., *Strukturális magyar nyelvtan 3. Morfológia*. Budapest, Akadémiai Kiadó. 75–136.
- KESZLER BORBÁLA 2000a. A névszótövek. In: KESZLER BORBÁLA szerk., *Magyar grammatika*. Budapest, Nemzeti Tankönyvkiadó. 175–182.
- KESZLER BORBÁLA 2000b. Az egyszerű és összetett mondat határsávja. In: KESZLER BORBÁLA szerk., *Magyar grammatika*. Budapest, Nemzeti Tankönyvkiadó. 461–471.
- KESZLER BORBÁLA 2000c. Az ige-tövek. In: KESZLER BORBÁLA szerk., *Magyar grammatika*. Budapest, Nemzeti Tankönyvkiadó. 95–103.

- KIEFER FERENC 1992. Bevezető. In: KIEFER FERENC szerk. *Strukturális magyar nyelvtan 1. Mondattan*. Budapest, Akadémiai Kiadó. 7–19.
- KIEFER FERENC 1999. Alaktan. In: É. KISS KATALIN–KIEFER FERENC–SIPTÁR PÉTER, *Új magyar nyelvtan*. Budapest, Osiris Kiadó. 187–290.
- KIEFER FERENC 2000. A morfológia. In: KIEFER FERENC szerk., *Strukturális magyar nyelvtan 3. Morfológia*. Budapest, Akadémiai Kiadó. 23–73.
- KIEFER FERENC–LADÁNYI MÁRIA 2000. A szóképzés. In: KIEFER FERENC szerk., *Strukturális magyar nyelvtan 3. Morfológia*. Budapest, Akadémiai Kiadó. 137–164.
- É. KISS KATALIN 1998. A generatív nyelvészet mint kognitív tudomány. In: PLÉH CSABA–GYÖRI MIKLÓS szerk., *A kognitív szemlélet és a nyelv kutatása*. Budapest, Pólya Kiadó. 23–39.
- É. KISS KATALIN–SZABOLCSI ANNA 1992. Grammatikaelméleti bevezető. In: KIEFER FERENC szerk., *Strukturális magyar nyelvtan 1. Mondattan*. Budapest, Akadémiai Kiadó. 21–77.
- KUHL, PATRICIA K.–STEVENS, ERICA–HAYASHI, AKIKO–DEGUCHI, TOSHISADA–KIRITANI, SHIGERU–IVERSON, PAUL 2006. Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science* 9/2: F13–F21.
- LACZKÓ KRISZTINA 2000. Az alaktan tárgya és alapkategóriái. In: KESZLER BORBÁLA szerk., *Magyar grammatika*. Budapest, Nemzeti Tankönyvkiadó. 37–50.
- LECANUET, JEAN-PIERRE 1998. Foetal responses to auditory and speech stimuli. In: SLATER, ALAN szerk., *Perceptual development. Visual, auditory, and speech perception in infancy*. Hove, Psychology Press. 317–355.
- LUKÁCS ÁGNES 2001. Szabályok és kivételek: a kettős modell érvényessége a magyarban. In: PLÉH CSABA–LUKÁCS ÁGNES szerk., *A magyar morfológia pszicholingvisztikája*. Budapest, Osiris Kiadó. 119–152.
- LUKÁCS ÁGNES 2005. A nyelvtani tudás gyökerei. In: GERVAIN JUDIT–KOVÁCS KRISTÓF–LUKÁCS ÁGNES–RACSMÁNY MIHÁLY szerk., *Az ezerarcú elme. Tanulmányok Pléh Csaba 60. születésnapjára*. Budapest, Akadémiai Kiadó. 38–49.
- LUKÁCS ÁGNES 2008. Nyelvtani feldolgozás. In: CSÉPE VALÉRIA–GYÖRI MIKLÓS–RAGÓ ANETT szerk., *Általános pszichológia 3. Nyelv, tudat, gondolkodás*. Budapest, Osiris Kiadó. 116–154.
- LUKÁCS ÁGNES 2014. Szótanulás. In: PLÉH CSABA–LUKÁCS ÁGNES szerk., *Pszicholingvisztika. Magyar pszicholingvisztikai kézikönyv 1–2*. Budapest, Akadémiai Kiadó. 521–549.
- LUKÁCS ÁGNES–PLÉH CSABA 2003. A nyelv idegrendszeri reprezentációja. In: PLÉH CSABA–KOVÁCS GYULA–GULYÁS BALÁZS szerk., *Kognitív idegtudomány*. Budapest, Osiris Kiadó. 528–560.
- LUKÁCS ÁGNES–GÁBOR BÁLINT–KEMÉNY FERENC–BABARCZY ANNA 2014. A nyelvtani kategóriák és szabályok elsajátítása: megfigyelési és kísérleti adatok. In: PLÉH CSABA–LUKÁCS ÁGNES szerk., *Pszicholingvisztika. Magyar pszicholingvisztikai kézikönyv 1–2*. Budapest, Akadémiai Kiadó. 483–520.
- MACWHINNEY, BRIAN 2003. A nyelvfejlődés epigenezise. In: PLÉH CSABA–KOVÁCS GYULA–GULYÁS BALÁZS szerk., *Kognitív idegtudomány*. Budapest, Osiris Kiadó. 505–527.

- MAJOROS MÁRIA 1998. Az anyanyelv mint a matematikai megismerést korlátozó modell. In: PLÉH CSABA–GYÖRI MIKLÓS szerk., *A kognitív szemlélet és a nyelv kutatása*. Budapest, Pólya Kiadó. 277–286.
- MAGYARI LILLA 2005. A nyelv miért nem olyan, mint a szem? In: GERVAIN JUDIT–KOVÁCS KRISTÓF–LUKÁCS ÁGNES–RACSMÁNY MIHÁLY szerk., *Az ezerarcú elme. Tanulmányok Pléh Csaba 60. születésnapjára*. Budapest, Akadémiai Kiadó. 452–460.
- MARCUS, GARY F. 1999a. Connectionism: with or without rules? Response to J. L. McClelland and D. C. Plaut. *Trends in Cognitive Sciences* 3: 168–170.
- MARCUS, GARY F. 1999b. Do infants learn grammar with algebra or statistics? Response to Seidenberg & Elman, Negishi, and Eimas. *Science* 284: 436–437.
- MARCUS, GARY F. 1999c. Reply to Seidenberg and Elman. *Trends in Cognitive Sciences* 3: 289.
- MARCUS, GARY F. 1999d. Reply to Christiansen and Curtin. *Trends in Cognitive Sciences* 3: 290–291.
- MARCUS, GARY F. 1999e. Rule learning by seven-month-old infants and neural networks. Response to Altmann and Dienes. *Science* 284: 875.
- MARCUS, GARY F.–VIJAYAN, SUJITH–BANDI RAO, SHOBA–VISHTON, PETER M. 1999. Rule-learning by seven-month-old infants. *Science* 283: 77–80.
- MCCLELLAND, JAMES L.–PLAUT, DAVID C. 1999. Does generalization in infant learning implicate abstract algebra-like rules? *Trends in Cognitive Sciences* 3: 166–168.
- MEHLER, JACQUES–JUSCZYK, PETER–LAMBERTZ, GHISLAINE–HALSTED, NILOFAR–BERTONCINI, JOSIANE–AMIEL-TISON, CLAUDINE 1988. A precursor of language acquisition in young infants. *Cognition* 29: 143–178.
- MÜLLER, RALPH-AXEL 1996. Innateness, autonomy, universality? Neurobiological approaches to language. *Behavior and Brain Sciences* 19: 611–675.
- NÁDASDY ÁDÁM–SIPTÁR PÉTER 1994. A magánhangzók. In: KIEFER FERENC szerk., *Strukturális magyar nyelvtan 2. Fonológia*. Budapest, Akadémiai Kiadó. 42–182.
- NEGISHI, MICHIO 1999. Do infants learn grammar with algebra or statistics? *Science* 284: 435.
- PAPP ISTVÁN 1966. *Leíró magyar hangtan*. Budapest, Tankönyvkiadó.
- PELUCCHI, BRUNA–HAY, JESSICA F.–SAFFRAN, JENNY R. 2009a. Learning in reverse: eight-month-old infants track backward transitional probabilities. *Cognition* 113: 244–247.
- PELUCCHI, BRUNA–HAY, JESSICA F.–SAFFRAN, JENNY R. 2009b. Statistical learning in a natural language by 8-month-old infants. *Child Development* 80: 674–685.
- PEÑA, MARCELLA–BONATTI, LUCA L.–NESPOR, MARINA–MEHLER, JACQUES 2002. Signal-driven computations in speech processing. *Science* 298: 604–607.
- PERRUCHET, PIERRE–PEEREMAN, RONALD 2004. The exploitation of distributional information in syllable processing. *Journal of Neurolinguistics* 17: 97–119.
- PINKER, STEVEN 1984. *Language learnability and language development*. Cambridge, Harvard University Press.
- PINKER, STEVEN 1994/2006. *A nyelvi ösztön. Hogyan hozza létre az elme a nyelvet?* Második, javított kiadás. Budapest, Typotex Kiadó.

- PLÉH CSABA 2000. A magyar morfológia pszicholingvisztikai aspektusai. In: KIEFER FERENC szerk., *Strukturális magyar nyelvtan 3. Morfológia*. Budapest, Akadémiai Kiadó. 951–1020.
- PLÉH CSABA 2006a. A gyermeknyelv. In: KIEFER FERENC főszerk., *Magyar nyelv*. Budapest, Akadémiai Kiadó. 753–782.
- PLÉH CSABA 2006b. Pszicholingvisztika. In: KIEFER FERENC főszerk., *Magyar nyelv*. Budapest, Akadémiai Kiadó. 725–752.
- PLÉH CSABA 2014. A pszicholingvisztika története. In: PLÉH CSABA–LUKÁCS ÁGNES szerk., *Pszicholingvisztika. Magyar pszicholingvisztikai kézikönyv 1–2*. Budapest, Akadémiai Kiadó. 3–116.
- PLÉH CSABA–LUKÁCS ÁGNES 2003. Nyelv, evolúció és az agy. In: PLÉH CSABA–KOVÁCS GYULA–GÜLYÁS BALÁZS szerk., *Kognitív idegtudomány*. Budapest, Osiris Kiadó. 485–504.
- PLÉH CSABA–LUKÁCS ÁGNES–KAS BENCE 2008. A szótár pszicholingvisztikája. In: KIEFER FERENC szerk., *Strukturális magyar nyelvtan 4. A szótár szerkezete*. Budapest, Akadémiai Kiadó. 789–852.
- PRÓSZÉKY GÁBOR 2000. A magyar morfológia számítógépes kezelése. In: KIEFER FERENC szerk., *Strukturális magyar nyelvtan 3. Morfológia*. Budapest, Akadémiai Kiadó. 1021–1063.
- REBRUS PÉTER 2000. Morfofonológiai jelenségek. In: KIEFER FERENC szerk., *Strukturális magyar nyelvtan 3. Morfológia*. Budapest, Akadémiai Kiadó. 763–947.
- RÓNA-TAS ANDRÁS 1978. *A nyelvrokonság. Kalandozások a történeti nyelvtudományban*. Budapest, Gondolat Kiadó.
- SAFFRAN, JENNY R.–ASLIN, RICHARD N.–NEWPORT, ELISSA L. 1996. Statistical learning by 8-month-old infants. *Science* 247: 1926–1928.
- SAFFRAN, JENNY R.–HAUSER, MARC–SEIBEL, REBECCA–KAPFHAMER, JOSHUA–TSAO, FRITZ–CUSHMAN, FIERY 2008. Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition* 107: 479–500.
- SAFFRAN, JENNY R.–NEWPORT, ELISSA L.–ASLIN, RICHARD N.–TUNICK, RACHEL A.–BARRUECO, SANDRA 1997. Incidental language learning: listening and learning out of the corner of your ear. *Psychological Science* 8: 101–105.
- SANDERS, LISA D.–NEWPORT, ELISSA L.–NEVILLE, HELEN J. 2002. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience* 5: 700–703.
- SANDERS, LISA D.–NEVILLE, HELEN J. 2003. An ERP study of continuous speech processing 1. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research* 15: 228–240.
- SANTELMANN, LYNN M.–JUSZYK, PETER W. 1998. Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition* 69: 105–134.
- SÁROSI ZSÓFIA 2005. Morfématörténet. In: KISS JENŐ–PUSZTAI FERENC szerk., *Magyar nyelvtörténet*. Budapest, Osiris Kiadó. 129–172, 352–371, 610–617, 719–724, 800–803.
- SAUSSURE, FERDINAND DE 1916/1997. *Bevezetés az általános nyelvészetbe*. Budapest, Corvina.

- SEIDENBERG, MARK S.–ELMAN, JEFFREY L. 1999a. Do infants learn grammar with algebra or statistics? *Science* 284: 434–435.
- SEIDENBERG, MARK S.–ELMAN, JEFFREY L. 1999b. Networks are not ‘hidden rules’. *Trends in Cognitive Sciences* 3: 288–289.
- SEIDENBERG, MARK S.–MACDONALD, MARYELLEN S.–SAFFRAN, JENNY R. 2002. Does grammar start where statistics stop? *Science* 298: 553–554.
- SIPTÁR PÉTER 1999. Hangtan. In: É. KISS KATALIN–KIEFER FERENC–SIPTÁR PÉTER, *Új magyar nyelvtan*. Budapest, Osiris Kiadó. 293–387.
- SOSA, ANNA V.–MACFARLANE, JAMES 2002. Evidence for frequency based constituents in the mental lexicon. Collocations involving the word *of*. *Brain and Language* 83: 227–236.
- SZILÁGYI N. SÁNDOR 1999. A szent mókus, avagy a módszer buktatói. In: POZSGAI PÉTER szerk., *Tűzcsiholó. Írások a 90 éves Lükő Gábor tiszteletére*. Budapest, Táton. 345–365.
- SZILÁGYI N. SÁNDOR 2004. *Elmélet és módszer a nyelvészetben — különös tekintettel a fonológiára*. Erdélyi Tudományos Füzetek 245. Kolozsvár, Erdélyi Múzeum-Egyesület.
- SZILÁGYI N. SÁNDOR 2013. A mi (rend)szertelen nyelvünk. Egy analógiás nyelvtan vázlat. In: KÁDÁR EDIT–SZILÁGYI N. SÁNDOR szerk., *Analógia és modern nyelvelírás*. Kolozsvár, Erdélyi Múzeum-Egyesület. 235–263.
- TAUZIN TIBOR M. 2010. A társas életmód hatása a mesterséges nyelvtan tanulására embernél és állatoknál. *Magyar Pszichológiai Szemle* 65: 389–412.
- THUMA ORSOLYA–PLÉH CSABA 2001. Kétértelműség és dekompozíció a magyar nyelvben. In: PLÉH CSABA–LUKÁCS ÁGNES szerk., *A magyar morfológia pszicholingvisztikája*. Budapest, Osiris Kiadó. 39–53.
- TURK-BROWNE, NICHOLAS B.–SCHOLL, BRIAN J.–CHUN, MARVIN M.–JOHNSON, MARCIA K. 2008. Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience* 21: 1934–1945.
- VELCSOV MÁRTONNÉ 1968. Alaktan. In: BENCÉDY JÓZSEF–FÁBIÁN PÁL–RÁCZ ENDRE–VELCSOV MÁRTONNÉ, *A mai magyar nyelv*. Budapest, Tankönyvkiadó. 85–113.
- WHITE, KATHERINE S.–PEPERKAMP, SHARON–KIRK, CECILIA–MORGAN, JAMES L. 2008. Rapid acquisition of phonological alternations by infants. *Cognition* 107: 238–265.
- WITTGENSTEIN, LUDWIG 1952/1992. *Filozófiai vizsgálódások*. Budapest, Atlantisz.