

14 ÉVES GYERMEKEK ÉLŐ NYELVHASZNÁLÁ- TÁNAK NÉHÁNY BELSŐ ÖSSZEFÜGGÉSE

ZSULÁN JULIANNA

Az itt következő adatok egy nagyobb empirikus vizsgálat részei; a teljes anyagból a kutatás céljaira és körülményeire, valamint az élőnyelvi töltelék- és hibajelenségekhez kapcsolódó részek korábban már megjelentek (Zsulán 1989–90); s ugyancsak napvilágot látott a korpuszban szereplő 155 tanulótól összegyűjtött kb. 114 ezer szöveg-
szó főneveinek jelentéstani vizsgálatáról szóló összegzés (Bicskeiné 1990). Így e közlemény nem tér ki egyetlen olyan témára sem, amely az említett írásokban megjelent.

Ahhoz azonban, hogy az alább következő elemzések világosak legyenek, mégis szükséges a hivatkozott első cikkből a korpusz jellegére vonatkozó közléseket megismételni.

Az itt közölt vizsgálat anyagát 155 nyolcadik osztályos tanulónak kb. 114 ezer szövegszónyi spontán beszélt nyelvi szövege adja. Őket egy nagyobb kutatási bázisból rétegezett mintavétellel választottuk ki úgy, hogy a kiválasztási szempontok az országos arányok leképezésére törekedtek az adatközlők neme, a szülők iskolai végzettsége és az iskola szociológiai környezete szempontjából.

Az adatgyűjtést megelőzően a kiválasztott tanulókkal csoportosan és egyénileg ismerkedő, gátlásoldó célú beszélgetést tartottunk. Ezt követően minden gyerekkel 2 x 5 percnyi magnetofonos felvétel készült két témáról: 1. egy történet elmondása (saját élmény, film vagy irodalmi mű cselekménye lehetett), 2. pályaválasztási elképzelések és motivációk, jövőkép. Az első téma tehát elbeszélő, a második értekező-fejtegető műfajú spontán szöveg alkotását kívánta; s vélhetőleg az első könnyebb, a második nehezebb feladatot rótt az adatközlőkre. A feladatokat előre nem, csak egyénenként, közvetlenül a felvétel előtt kapták meg, s a már elhangzott interjúk alanyai „kötelező titoktartást” fogadtak. Erre feltétlenül szükség volt a spontaneitás biztosítása végett. S mivel ezek a 14 éves gyerekek élvezték a játékos hangűtést, tapasztalataink szerint többnyire sikerült is a „titoktartást” megőrizni.

A felvétel előtt folyamatos beszédre kértük őket, segítséget csak akkor adtunk, ha az adott időn belül kifogytak a mondanivalóból, s akkor is csak olyan segítséget, amelynek a megfogalmazása nem adott nekik új kifejezéseket, ötleteket a folytatásához (pl. egyáltalán nem tettünk fel teljes formájú kérdéseket, sem kiegészítendőket, sem eldöntendőket).

Az elhangzott szövegeket ellenőrzés után számítógépre vittük, majd Word-program segítségével részint kódoltuk a vizsgált nyelvi jelenségek szerint, részint egyénenkénti szógyakorisági listákat készítettünk (ezekből készült a teljes korpusz gyakorisági szótára). A kódjelek összesítése és a statisztikai adatok kiszámítása szintén

számítógépen történt. Megjegyzendő még, hogy – mivel a minta kiválasztását külső személy végezte – a kódolás időszakában az adatközlőkről mindössze annyit tudtunk, melyik iskolába járnak, sem tanulmányi eredményük, sem társadalmi háttérük nem volt ismeretes számunkra; így a szövegek egyes nyelvi jelenségeinek megítélésében még akaratlanul sem lehettünk részrehajlók.

Ebben a dolgozatban a nyelvhasználati jelenségek közötti összefüggések feltárására a statisztikai módszerek közül a korrelációs számítást használtuk fel. Ez a módszer a nem oksági, valószínűségi kapcsolatok vizsgálatára alkalmas, hiszen képes meghatározni a sztochasztikus kapcsolatok szorosságának mérőszámát: a korrelációs együtthatót. Az alábbiakban az összes elemzés ezzel az adattal dolgozik. A korrelációs együttható (a dolgozat szövegében következetesen: „a korreláció”) értelmezése a következő:

- ha értéke:
- 0, akkor a két jelenség között nincs kapcsolat;
 - 1, akkor a két jelenség között igen erős fordított arányosság áll fenn;
 - +1, akkor a két jelenség között igen erős egyenes arányosság áll fenn;
 - 0 és 1 közé esik (akár negatív, akár pozitív előjellel), akkor
 - 0,0 – 0,3 gyenge összefüggést,
 - 0,3 – 0,6 mérsékelt,
 - 0,6 – 0,9 erős,
 - 0,9 – 1,0 igen erős kapcsolatot jelez.

A korpusz jellegének és a statisztikai eljárásnak az ismertetése után álljon itt a dolgozatban használt nyelvészeti terminológiára vonatkozó néhány megállapítás is! Mivel a feldolgozott anyag spontán beszélnyelvi szöveg, a dolgozatban előforduló „mondat”, „mondategység” kifejezések (amelyek az írott szövegre vonatkozó terminus technikusok) „megnyilatkozás”, „nyilatkozat”, illetve „egység”, „egész” értelemben használatnak. Továbbá: az „élőnyelv”, „beszélt nyelv”, „elhangzó szöveg”, „élőszóban elhangzó mondat” kifejezések mind a „spontán beszélnyelvi szöveg” szinonímáiként kerültek a dolgozatba, bár kétségtelen, hogy jelentésük a szakirodalomban nem teljesen egyezik az itt használttal.

E dolgozatban a következő nyelvhasználati jelenségek közötti belső összefüggések matematikai-statisztikai vizsgálatának eredményei kaptak helyet:

a) A korpuszban előforduló nyelvi jelek és jeltípusok közötti arányok mérésére az *iteráció* ($J : J_i$) és a *szókincsgazdagság* ($J_i : \sqrt{J}$) mutatóját használtuk fel. E két mutató a magyar szakirodalomból is jól ismert (Nagy 1980, 25., Zsilka 1974, 22.); köztudott az is, hogy az utóbbi csak 2000 szövegszó feletti korpuszban ad torzítatlan képet. Hasonló a gond az iterációs mutatóval is. Minél rövidebb ugyanis egy szöveg, annál kisebb a szóismétlések (nyelvhasználati hibák!) valószínűsége, s a szóismétlődés-ké (a nyelv természetes jelenségei, pl. a névelők, névmások, vagy a téma részletezésé-

hez feltétlenül szükséges szavak esetében) is meglehetősen csekély. A hosszabb szövegekben az ismétlődés száma törvényszerűen nő, az ismétlés csak fogalmazási ügyetlenség miatt. Ha pedig azonos terjedelmű két szövegből az egyik részletesen kifejti gondolatait, a másik csak felvillantja őket, az előbbiben biztosan több lesz az ismétlés és ismétlődés (a feltétlenül szükséges szavak miatt meg esetleg a fogalmazási gyakorlatlanság miatt is), vagyis: az első szöveg iterációs mutatója kedvezőtlenebb lesz, mint a másodiké, pedig a szövegminőség az első esetben jobb. Ezért az iterációs adatot ebben a korpuszban kétféle módon értelmeztük: elvileg akkor kedvező a közleményben a jellettípus viszony, ha a mutató értéke alacsony; de a fejletlen vagy fejlődőben levő nyelvhasználatú szövegek alacsony értékű mutatói éppen a rövidség és kifejtetlenség miatt alakultak ki, ez esetben tehát a magasabb értékek a kedvezőbbek (tehát ha a közlemény hosszabb és kifejtettebb). Az itt vázolt problémák megkérdőjelezzik e mutatók használatát; s hogy mégis felhasználtuk őket, annak oka az a feltevés volt, hogy a várható torzítások és értelmezési nehézségek ellenére is képesek lesznek a többi mutatóval együtt jellemezni az egyéni nyelvhasználatot.

E két mutatónak egymással és a többi mutatóval kialakult kapcsolatait az 1.1.

– 1.4. pontban tárgyaljuk.

b) Mutatót képezünk a szövegben előforduló *elvont jelentésű jellettípusok* arányának vizsgálatára, s ennek segítségével az egyéni szókészlet absztrakciós szintjére nézve kívántunk adatokat szerezni. (Az osztályozás szempontjairól és módszereiről l. Nyr. 114: 68-9.) Ezen mutató kapcsolatainak elemzése a 2.1. – 2.3. pontban található.

c) A terjedelem mérésére a *teljes korpusz*, a *tartalmi egységek* és a *mondategységek szövegszavainak számát* használtuk. Tartalmi egységnek azokat a szövegegységeket neveztük, amelyek neve az írásos közleményekben bekezdés; mivel azonban ezek elkülönítése az elhangzó beszédműben még a mondategységek határainak megállapításánál is nehezebb (és ezért szubjektívebb is), célszerűnek látszott másik elnevezés keresése. A tartalmi egységek határait - hasonlóan a drámai művek felépítésében a jelenetek tagolásának elvéhez - ott húztuk meg, ahol vagy az előbbiekhöz képest teljesen új gondolat kezdődött, vagy pedig olyan motívummal bővült a szöveg, amely újabb kifejtést igényelt. Ezek az új motívumok a szövegben mindig rémák, de természetesen nem minden rémához tartozik új tartalmi egység, csak azokhoz, amelyek új gondolatsort indítanak meg. Ebben a korpuszban a tartalmi egységek terjedelme igen változó: lehet bekezdésnyi vagy csupán egy-két mondategységnyi is.

A terjedelemre vonatkozó mutatók egymással és a többi mutatóval kialakult kapcsolatait a 3., 3.1. – 3.3. pontban olvashatók.

d) A gondolatok grammatikai strukturáltságának (alárendeltségi fokának) vizsgálatában három mutatót alkalmaztunk: 1. a *mondategységek* és a *mondatrészek átlagos mélységének* mutatóját, (ezek lényegében súlyozott átlagszámítással jönnek

létre az egyes szintek népességének a szint sorszámával való megszorzása útján), valamint a két valódi mélységmutató összevonásából nyert *átlagos mélységmutatót*.

Ezen mutatók kapcsolatainak elemzését a 4. pont végzi el.

Az alább közlendő eredmények segítségével nemcsak a fejlődőben levő élő nyelvhasználat néhány jellegzetességére kívánunk rámutatni, hanem arra is, hogy a felsorolt mutatók hogyan működnek, mennyire adnak jellemző képet a vizsgált nyelvi jelenségekről, s alkalmasak-e későbbi felhasználás esetén az egyéni nyelvhasználat (itt az elhangzó spontán beszéd) jellemzésére.

A statisztikai úton nyert adatok elemzésének sorrendje végig azonos: először az adott csoporton belüli szignifikáns összefüggések értelmezését végezzük el, ezt követik a nem szignifikáns adatokra vonatkozó megállapítások.

1. A jel-jeltípus viszony két mutatójának kapcsolatai

1.1. Az *iteráció* és a *szókincsgazdagság* mutatója egymással 0.2197 értékű, a mutatók elvi irányai miatt fordított arányú, gyenge összefüggésben van. A fordított arány a két mutató kiszámításának módját ismerve érthető, hiszen egymásnak lényegében reciprokai.

A gyenge arányú összefüggés azonban elgondolkodtató, hiszen a tényezők azonosságából az következne, hogy ennél lényegesen erősebb legyen a két mutató közötti korreláció. E kapcsolatgyengülés okául felvetődhet a mutatók elégtelen működése és/vagy érvényességük korlátozottsága. E kérdést a szakirodalmi utalással együtt a bevezetőben már említettük; de hogy még e korlátok ellenére is van jellemző erejük, azt az alábbiakkal bizonyítva látjuk.

Ha ugyanis összehasonlítjuk a két - kétségtelenül vitatható - mutató kapcsolatrendszerét, láthatjuk, hogy szignifikáns összefüggéseik száma közel azonos, s ezek felerészben meg is egyeznek. (A jelölések: *xx erősebb, x gyengébb* korrelációt jelent.)

A mutató neve	korrelációja	
	az iterációval	a szókincsgazdagsággal
elvont jeltípusok aránya	x	xx
a korpusz szövegszavai	xx	xx
a tart.egys. "	xx	xx
a mondategys. "	xx	x
a mondategys. mélysége	xx	xx
a mondatrészek "	o	xx
összevont mélységmutató	xx	xx
'hezitáció	o	x

modorosság	x	0
a kötőszók hibái	xx	0
összes hiba	x	0
összes hiba és hiány	x	0

* E mutatók tárgyalása nem e dolgozat keretébe tartozik (vö. NrNyt. XXXIII. 91-103.), a teljesség kedvéért azonban álljanak itt ezen adatok is.

Vagyis: a korrelációk

megegyeznek

eltérnek

xx - xx módon 4 esetben

xx - 0 módon 2 esetben

x - xx módon 2 esetben

x - 0 módon 4 esetben.

Ez az adatsor azt látszik bizonyítani, hogy az egyéni beszédképesség sokoldalú jellemzésében mindkét mutatót (együtt!) haszonnal alkalmazhatjuk. További részletes elemzésük szintén ezt kívánja bizonyítani.

1.2. A jel-jeltípus viszony összefüggése az elvont jeltípusok arányával

A korrelációs adatok a következők: az elvont jeltípusok aránya az iterációval 0.2438, a szókinccsgazdagsággal 0.3996 értéket mutat, vagyis az előbbi a két jelenség közötti gyenge, a második mérsékelten erős kapcsolatra utal. A két jel-jeltípus mutató reciprok mivolta szerint a korrelációs adatok ellentétes jelentést hordoznak: az elvont jelek aránya az iterációval fordított, a szókinccsgazdagsággal pedig egyenes arányosságban áll. Eszerint az elvont jelentésű jeltípusok arányának szövegbeli növekedése maga után von(hat)ja az iterációs tendencia csökkenését, illetve a szókinccsgazdagság növekedését. Így azt mondhatjuk, hogy a gazdagabb szókinccs, a választékosabb szóhasználat százalékos arányban kifejezhetően több elvont jelentésű jeltípust képes alkalmazni, mint a gyengébben fejlett (vö. Nyr. 114: 73-85.); illetve: magasabb absztrakciós szinten valószínűleg kisebb mértékű a fogalom-, viszony- és határozószók ismétlési tendenciája a szövegben.

1.3. A jel-jeltípus arány és a szövegegységek szószáma közötti kapcsolat

A teljes szöveg szószáma azért jellemző adat, mert adott idő (10 perc) alatt megalkotott szöveg hosszát mutatja, vagyis azt, hogy melyik adatközlő hány szövegszóból volt képes mondandóját összeállítani ennyi idő alatt azonos témáról, műfajban és teljesen azonos közlési körülmények között (arra törekedtünk pl., hogy még a segítő közbevetések, biztatások is azonosak legyenek).

A tartalmi egységek hossza igen változó; vannak olyan korpuszok, amelyek sok témát érintenek részletezés nélkül, és vannak olyanok is, amelyek részletesen kifejtik gondolataikat. A tartalmi egység szószáma sem nem szűkíthető, sem nem bővíthető a végtelenségig: a sűrítésnek a minimálisan kötelező információmennyiség kényszere szab határt, a bővítésnek pedig a redundancia; s általában a különféle témák, kommunikációs helyzetek és műfajok többé-kevésbé meghatározzák az optimális szószámot.

A mondategység szószáma tekintetében lényegesen kevesebb a variációs lehetőség, mint az előző két esetben; e korpuszban ez az adat az ilyen életkorú adatközlők ilyen műfajú és témájú szövegeiben tipikusan jellemzőnek tűnik; s általánosabban véve is igaz az, hogy a mondategység szószáma a magyar nyelv balra bővülő szerkesztésmódja következtében nem lehet túlságosan nagy, különösen az elhangzó beszédben, amelyben a sokszorosan bővített és alászerkesztett mondat megértésének kommunikációs akadályai vannak.

A fenti megállapítások ismeretében lássuk a korrelációs adatokat:

<u>Szószám</u>	<u>iteráció</u>	<u>szókincsgazdagság</u>
a teljes szövegben	0,7641	0,7616
a tart. egységekben	0,4843	0,5266
a mondategységekben	0,2718	0,2087

A három adatsor közül az első, a jel-jeltípus viszonyra és a teljes szöveg terjedelmére vonatkozó adatok a legmagasabbak és rendkívül közeli értékűek: az iterációval erős fordított, a szókincsgazdagsággal pedig erős egyenes arányosság alakult ki. Ez nyilvánvaló: a gazdagabb szókincs, amelynek birtokában a szövegben kisebb a szóismétlések valószínűsége, nagyobb szövegterjedelmet tesz lehetővé.

A második adatsor a jel-jeltípus arány és a tartalmi egységek szószáma közötti mérsékelt erősséget mutatja. A korrelációs adatnak a teljes szöveghez képest való csökkenése minden valószínűség szerint arra utal, hogy - bár az előző bekezdésben a szókincs, a szóismétlések és a terjedelem összefüggéséről tett megállapítás itt is igaz - ezen összefüggés érvénye a szövegnél alacsonyabb szintű egységben korlátozódik, s nyilvánvalóan éppen a téma által meghatározott módon. Am fel kell figyelni arra, hogy a két korreláció közül a szókincsgazdagságra vonatkozó a nagyobb értékű, s ez minden bizonnyal azt jelenti, hogy az egyéni szókincs terjedelme erősebben befolyásolja a tartalmi egységek szószámának alakulását, mint a szóismétlési/ismétlődési tendencia.

A harmadik adatsorban látható adatok a vizsgált jelenségek között gyenge összefüggéseket jelölnek: a mondategység szószáma fordítottan arányos az iterációval és egyenesen a szókincsgazdagság mutatójával. Eszerint azokban a szövegekben, amelyekben az átlagnál kevesebb szóból épülnek fel a mondategységek, előfordulhat, hogy erősebb a szóismétlési tendencia; illetve: a gazdagabb szókészlet nagyobb szószá-

mú mondat egységeket szerkesztésére tesz képessé. Az utóbbi megállapítás evidencia; az előbbi viszont talán új nézőpontot jelenthet a fejletlen-fejlődőben levő nyelvhasználat vizsgálatához. Mindemellett a korrelációs adatok gyengesége arra is utal, hogy a mondat egységek szószámát ebben a vizsgálatban nem igazán befolyásolta az egyéni szóképlet nagysága, hanem inkább a közlési körülményeket jellemző sajátásnak bizonyult.

Mindezeken kívül a fenti adatsor fényt vet az iterációs és szókincs gazdagsági mutatók működésére is. A teljes szöveg terjedelmére vonatkozó adatpár azt látszik bizonyítani, hogy a jelek és jeltípusok aránya mindkét mutatóval közel azonos megbízhatósággal mérhető; de azt is bizonyítja, hogy torzításuk nem abszolút mértékű, hiszen akkor a korrelációjuk is magasabb értékű lenne. Így ha azoknak a teljes szövegre vonatkozó nagysága nem 0,9, hanem csak 0,7, ez mindenképpen azt jelzi, hogy a képletek – ha nem is tökéletesen – képesek a különbségeket regisztrálni a különböző hosszúságú szövegek szóképletének terjedelmére nézve.

A tartalmi egységekre vonatkozó két korreláció különböző nagysága arra enged következtetni, hogy az azonos jelenséget mérő különböző mutatók közül a szókincs gazdagság adja a szorosabb összefüggést, valószínűleg azért, mert mérési módszere kevésbé torzít, mint az iterációé.

1.4. A jel-jeltípus arány és a mélységmutatók kapcsolata

A mondat egységek és a mondatrészek mélysége, valamint az e kettőből összevont mélységmutató - egy kivétellel - szignifikáns összefüggésben van az iterációval és a szókincs gazdagsággal. Korrelációs adatai:

	<u>iteráció</u>	<u>szókincsgazdagság</u>
a mondat egységek mélysége	0,4565	0,2738
a mondatrészek mélysége	0,0013	0,2868
	(nem szign.!)	
összevont mélység	0,4281	0,3737

Először vegyük szemügyre az iterációnak az első két mélységmutatóval kialakult kapcsolatait! Az első adat mérsékelten erős, az iteráció elvi iránya miatt fordított arányosságot jelez, s ez arra utal, hogy a mondat egységek mélységének csökkenésével, vagyis erőteljesebb mellérendelő szerkesztésmóddal a szóismétlések száma növekszik, az alárendelő mondatok arányának növekedésével pedig csökken. Ez az összefüggés nyilvánvaló, hiszen az összetett mondatok szerkesztésében az alárendelés inkább kizárja a fogalomszók ismétlését, mint a mellérendelés, amelyben az ilyen szavak - főleg fogalmazási gyakorlatlanság esetén - gyakrabban felbukkanhatnak. Az első oszlop második adata érdekes megfigyelésre ad alkalmat: az iteráció, amely a mondat egységek

mélységével mérsékelten erős kapcsolatban van, a mondatrészek mélységével egyáltalán nem függ össze (a korrelációs adat majdnem $\emptyset!$), s ez annak bizonyítéka, hogy a szóismétlési tendencia a mondatrészek hierarchiájából szinte teljesen hiányzik.

Ha most a második oszlopot nézzük, látható, hogy a szókincsgazdagság mutatója megközelítőleg azonos erősségű kapcsolatot jelez a mondategységek és a mondatrészek mélységével. Ezek a gyenge, egyenes arányú összefüggések arra utalnak, hogy az egyéni szókészlet terjedelmének növekedésével együtt jár mind a mondategységek, mind a mondatrészek erőteljesebb megszerkesztettsége, azaz az alárendelő konstrukciók növekedő száma. Az a beszélő tehát, akinek birtokában megfelelő nagyságú és összetételű szókincs van, nyelvi tevékenységében képes a dolgok és jelenségek összefüggéseit nemcsak tartalmi-logikai (mellérendelő), hanem grammatikai (alárendelő) struktúrákban megfogalmazni.

A szókincs terjedelmére és a mélységre vonatkozó korrelációs adatok közül a legmagasabb értékek a táblázat harmadik sorában találhatók: ezek az összevont mélységmutató mérsékelten erős kapcsolatát mutatják mind az iterációval, mind a szókincsgazdagsággal, és ez az előbbi esetben fordított, az utóbbiban egyenes arányosságot jelent. Eszerint a szövegben a mondategységek és a mondatrészek alárendelő struktúráinak nagyobb aránya együtt jár a szóismétlések számának csökkenésével; vagy a másik mutató felől megközelítve: a szókészlet terjedelmének növekedése maga után vonja mindkét nyelvi szinten a mélyebb szerkezetek arányának növekedését.

Annak, hogy a három mélységmutató közül éppen ez utóbbi összefüggés a legerősebb, értelmezésünk szerint a vizsgálati módszerre utaló jelentősége van: bizonyítja, hogy érdemes volt ezzel az összevont adattal kísérletezni, hiszen képes arra, hogy jellemző és összefüggéseiben is helytálló képet adjon a mélységi viszonyokról.

2. Az elvont jelentésű jeltípusok arányának kapcsolatai

2.1. Összefüggés a jel-jeltípus arányokkal

A kérdést az 1.2. pontban részletesen tárgyaltuk, itt - a teljesség kedvéért - csupán a fő következtetést ismétljük meg. Eszerint a fejlődő nyelvhasználatú élőnyelvi szövegekben az elvont jelentésű jeltípusok arányának növekedése együtt jár a szóismétlések ritkábbá válásával (vagyis a magasabb absztrakciós szókincs segítségével a 14 éves beszélő inkább el tudja kerülni a szavak ismételtetésében megnyilvánuló fogalmazási gyakorlatlanságot, mint az a társa, akinek szókészletében kisebb arányúak az ilyen szavak); illetve: az egyéni szókincs gazdagodása, a nagyobb lexikai választék több elvont jelentésű szó használatát valószínűsíti.

2.2. Az elvont jelentésű jeltípusok aránya és a szövegegységek terjedelme közötti viszonyok

Az elvont jelentésű jeltípusok aránya a teljes szöveg terjedelmével 0,3897, a tartalmi egységével 0,5942, a mondategységével 0,1025 értékű korrelációt mutat. Ebből az adatsorból világosan kitűnik, hogy a felhasznált szókészlet elvontságának mértéke szignifikáns kapcsolatban van az adott idő alatt létrehozott teljes szövegben és annak tartalmi egységeiben levő szavak számával, de nincs kapcsolatban a mondategység szószámával. Ez utolsó adat ismét a fentebb már érintett megállapítást támasztja alá: ebben az életkorban, ilyen közlési körülmények között, adott műfajok és témák követelménye szerint a mondategységen belüli szószám alig jellemző az egyénre, inkább általános nyelvhasználati sajátásnak tekinthető; ezért szinte teljesen független a beszédtevékenység egyéb elemeitől, így a szókészlet minőségi összetételétől. Ugyanezt támasztja alá egyébként a mondategység szószámáról nyert alapstatisztikai adatsor is (átlag: 5,12, modus: 5,14, relatív szórás 15,44%, a megoszlási sorban az átlag alatti és feletti, legközelebbi osztályközökben található a teljes mintának 75%-a).

Ezután térjünk vissza a szignifikáns összefüggésekre! Az adott idő alatt létrehozott teljes szöveg és a tartalmi egységek szószáma egyaránt mérsékelten erős, egyenes arányú összefüggést mutat az elvont jelentésű jeltípusok mennyiségével, de úgy, hogy ez a kapcsolat lényegesen erősebb a tartalmi egységek szószáma esetében. Ezt figyelemre méltó adatként kell értékelnünk, mivel azt jelzi, hogy a kisebb szövegegységek kifejtésében az elvont jelentésű jelek szerepe nagyobb, mint a teljes szövegében. Mit jelent ez? Azt, hogy egy-egy gondolati egységet akkor lehet bővebben kifejtteni, ha a beszélő a szövegépítésben nem csupán az érzéki úton nyert képzeteit szimbolizáló jeleket használja, nemcsak egymás mellé állítja a dolgokról és jelenségekről szóló mondandójának egyes elemeit, hanem összefüggéseket is keres köztük, azaz képes a viszonylatokat, relációkat kifejező szavak használatára is, pl. okokat, következményeket keres, hasonlít, magasabb halmazokba sorol, belső, lelki, emberi kapcsolatokat fejez ki, érvel vagy cáfol stb.; vagyis ha egy-egy fogalomszó szemantikai mezőösszefüggéseit minél szélesebb körben megnevezi. Persze így a tartalmi egység nemcsak több szóból áll, hanem valószínűleg megszerkesztettebb, mélyebb konstrukció lesz, tehát a mélységmutatók is növekednek (erről 1. még a 2.3. pontban).

A vizsgált jelenségek közötti, megközelítőleg 0,6 értékű korreláció azonban azt is jelzi, hogy ez az összefüggés a korpuszban nem érvényesül általánosan. A feldolgozott szövegek tanúsága szerint ugyanis (a várakozásnak megfelelően) az elbeszélő műfajú szövegekben általában alacsonyabbak az elvontsági értékek, mint az értekező-fejtegető műfajúakban. Ez evidencia, hiszen az események elmondása közben a beszélő jó és kifejtett szöveget hozhat létre akkor is, ha többségében konkrét jelentésű, érzéki tapasztalatokat leképező szavakat alkalmaz. Eszerint az egyéni szóbeli kifejezőképesség

mérésére mindenképpen legalább a fent említett két műfaj együttes alkalmazása célszerű, különösen a gyakorlatlan, fejletlen/fejlődőben levő nyelvhasználat esetében, hiszen a fejlett nyelvhasználatú művelt felnőtt nemcsak fejtegetéseiben tud absztrakt szintre emelkedni, hanem képes arra is, hogy az események elmondása közben elszakadjon az érzékelhető valóság kizárólagos megfogalmazásától; a fejletlen szövegalkotás viszont éppen ebbe kapaszkodik.

E kérdéshez érdekes adalékul szolgálhat a vizsgálatnak az a tapasztalata, hogy az elbeszélő szövegcsoporton belül igen nagy különbségek mutatkoztak az elvontság tekintetében aszerint, hogy milyen eredetű volt az az esemény, amit az adatközlő elmondott. Az olvasmányélményt ismertető szövegekben lényegesen magasabb elvontsági arány tapasztalható, mint azokban, amelyek egy film cselekményét mondták el. Ez utóbbiak - főleg a leggyengébbek - csupán a látványt regisztrálták, az előbbiek többnyire képesek voltak összefüggések, viszonylatok, belső tartalmak kifejezésére is (még a gyengébb képességűek is!). Persze az erre szolgáló nyelvi jeleket megtanulhatták magából az olvasmányból is (és ez az egyéni szókészlet fejlődésének egyik legbiztosabb útja); a film azonban - még ha nem kommersz, amilyenekről ezek a szövegek beszélnek, hanem művészi, akkor sem - nem tudja szavakban kifejezni képi mondandóját, a szókincsfejlesztésben ezért közvetlen hatásával nem számolhatunk.

2.3. A szókincs elvontságának összefüggése a mondategységek és a mondatrészek mélységével

Az előző pontban utalás történt arra, hogy azok a szövegek, amelyekben az elvont jelentésű jeltípusok aránya nagyobb, képesek az összefüggéseket is érzékeltetni, s ezzel tartalmi egységeiket részletesebben kifejteni, ez pedig mélyebb szerkezeteket hoz létre mind a mondategységek, mint a mondatrészek szintjén. Ez a megállapítás a korrelációs adatokkal bizonyítható is: az elvont jeltípusok kapcsolata a mondategységek mélységével 0,3060, a mondatrészekével 0,3027, az összevont mélységmutatóval 0,4110 értékű, egyenes arányú összefüggést mutat. Ezen adatok szerint tehát a magasabb absztrakciós szintet reprezentáló jeltípusok nagyobb szövegbeli aránya és a mondategységek/mondatrészek mélysége között mérsékelten erős összefüggés áll fenn; s ez, összekapcsolva a szókészlet terjedelmének és minőségének - már tárgyalt - szignifikáns viszonyával, azt bizonyítja, hogy az egyéni szókincs nagysága és minőségi összetétele, valamint az alárendelő mondategységek és mondatrészek beszerkesztésében alkalmazott struktúrák kölcsönösen befolyásolják egymást, s fejlettségük együttesen segíti elő a kifejezendő tudattartalom adekvát nyelvi megfogalmazását.

E két adat között nincs lényeges nagyságrendi különbség, a csekély eltérés mégis azt sejteti, hogy a szókincs elvontságának mértéke a mondategység alárendeltségi fokát valamivel erőteljesebb befolyásolja, mint a mondatrészeként.

3. A szövegegységek terjedelmének kapcsolatai

A három szövegegység szószámának egymással kialakult kapcsolatai a következők:

a teljes szöveg és a tartalmi egységek között:	0,6429,
a teljes szöveg és a mondategységek között:	0,3032,
a tartalmi egységek és a mondategységek között:	0,2669.

Az első sor adata erős, egyenes arányú összefüggést jelez a teljes szöveg és a tartalmi egységek hossza között, és ez teljesen nyilvánvalónak látszik, hiszen ha valaki képes részletesen kifejteni a teljes témát, az képes ugyanerre a tartalmi egységekben is. Ez a vizsgált korpusz hosszú szövegeinek nagyobb részében így is van, de némelyekben éppen ennek ellenkezője tapasztalható. Ezekben a beszélők sok, egymásra torlódó közlendőjüket inkább csak jelzik, de nem fejtik ki őket, egyik gondolatról a másikra csaponganak (különösen a vizuális élményről szóló elbeszélésekben). E két szélsőség a tárgyalt korreláció nagyságrendjét bizonyára erősen befolyásolja, hiszen ha a „hosszú szöveg - hosszú tartalmi egységek” összefüggés valóban evidencia lenne, a korrelációs adat feltétlenül magasabb lenne a kapott 0,6 értéknél.

A mondategység szószáma kapcsolatba hozható a teljes szöveg és a tartalmi egységek terjedelmével is, ám ez a két korrelációs adat e szövegegységek között lényegesen gyengébb az előbb tárgyaltnál, csupán mérsékelt, illetve gyenge, egyenes arányú összefüggést mutat. Eszerint a teljes szöveg és a tartalmi egységek nagyobb terjedelme és a mondategységbeli szószám emelkedése együtt jár; vagy fordítva: az az egyén, aki képes több szóból felépíteni mondategységeit, képes lehet bővebb kifejtésre a magasabb nyelvi szinteken is. Ez a tendencia azonban az adatok szerint csak gyengén érvényesül, s ennek oka bizonyára - különösen a fejlődő nyelvhasználatú gyermekek elhangzó beszédében - nemcsak a nyelvünk szerkesztésmódjával összefüggő megértésbeli akadály, hanem a még nem kifogástalanul működő mondat szerkesztési szabályrendszer. Ha ezek gátat vetnek a mondat szószáma emelkedésének, az új információk új mondategységekbe kerülnek; azaz a tartalmi egységek és a teljes szöveg szószámának emelkedése inkább a bennük lévő mondategységek számának emelkedésével jár, mint a mondategységeken belüli szószám lényeges növekedésével. E nyelvi szinten tehát az adatok szerint ebben az életkorban (még) a mondat bővíthetőségének viszonylag kevés az egyéni variációs lehetősége.

3.1. A szövegegységek terjedelmének összefüggése a jel-jeltípus aránnyal

Ezt a kérdést fentebb az 1.3. pontban részletesen tárgyaltuk. A szöveg/tartalmi egység és a jel-jeltípus arány közötti összefüggésről szóló fő következtetésünk ott az a nyilvánvaló igazság volt, hogy a gazdagabb szókincs nagyobb szövegterjedelmet tesz

lehetővé, és csökkenti azon szóismétlések számát, amelyek fogalmazási hibaként kerültek a szövegbe. Megjegyzendő azonban, hogy az iterációs mutató értelmezésének másik lehetősége szerint (erre a bevezetésben utaltunk) ez az összefüggés úgy is felfogható, hogy a gazdagabb szókincs nagyobb szövegterjedelmet tesz ugyan lehetővé, ám ez a fejlődőben lévő nyelvhasználat esetében nagyobb arányú szóismétléssel is együtt jár(hat), ha az adatközlő még nem tud élni a nyelvi szinonímia lehetőségeivel.

A mondategységek szószámát a szóismétlések alig befolyásolják; tehát ez a jelenség minden bizonnyal nem a mondategység, hanem az ennél magasabb szövegegységek szintjén ható nyelvhasználati tény.

Az e körbe tartozó összefüggések közül a leggyengébb, még szignifikáns korreláció a szókincsgazdagság mutatója és a mondategységek szószáma között mutatkozik, s ez azt jelenti, hogy a szövegépítésben felhasznált egyéni szókészlet terjedelme nem befolyásolta lényeges módon azt, hogy ki hány szóból építette fel a mondategységeit. Valószínű, hogy ez az összefüggés az élőnyelvi mondatalkotásra fokozottan érvényes.

3.2. A szövegegységek terjedelmének összefüggése az elvont jelentésű jeltípusok arányával

Ezt a kérdést a 2.2. pontban tárgyaltuk. Eszerint az elvont jelentésű jeltípusoknak a teljes szöveg/tartalmi egységek terjedelmével mutatott összefüggése azt jelenti, hogy egy-egy gondolatsor bővebb kifejtése akkor lehetséges inkább, ha a beszélő a szövegszerkesztésben nemcsak az érzéki úton nyert képzeteit szimbolizáló jeleket használja, hanem képes magasabb absztrakciós szint megragadására is. Abból, hogy a mondategység szószáma nincs szignifikáns összefüggésben az elvontság mutatójával, arra következtettünk, hogy az ilyen korú adatközlők elhangzó beszédműveiben a mondategység terjedelmét a szókincs minőségi összetétele alig befolyásolja.

3.3. A szövegegység terjedelmének összefüggése a mélységmutatókkal

E mutatók kettő kivételével szignifikáns összefüggésben vannak egymással:

		mélység		
		<u>mondategység</u>	<u>mondatrész</u>	<u>összevont</u>
szó- szám	teljes szöveg	0,4831	0,1571	0,5170
	tart. egység	0,3815	0,2834	0,4728
	mondategység	0,0975	0,3445	0,2347

Az adatokból első közelítésben az olvasható le, hogy a teljes szöveg és a tartalmi egység terjedelmének a mélységmutatókkal való egyenes arányossága tendenciájában megegyezik: legerősebb az összefüggésük az összevont mutatóval, ennél gyengébb a mondategységek mélységével, leggyengébb a mondatrészek mélységével, olyannyira,

hogy a teljes szöveg – mondatrészek viszonylatában ez már a szignifikancia alsó határa alá esik.

Másféle a mondategység terjedelmének kapcsolatrendszere: a mondatrészek mélységével van a legszorosabb, az összevont mutatóval ennél gyengébb kapcsolatban; a mondategységek mélységével kialakult, igen alacsony értékű korrelációs adat azonban szinte kizárja köztük az összefüggést.

Akik a nyelvhasználatot valamilyen szinten vizsgálják, egyetértenek abban, hogy a mondategységek és mondarészek erőteljesebb alárendelő szerkesztésmódja a bővebb terjedelemmel és a minőségileg – mennyiségileg fejlettebb szókészlettel együtt a fejlettebb nyelvhasználat jele. E tényezők közül a szókészletre és a terjedelemre vonatkozókról az eddigiekben már megállapítottuk, hogy valóban szorosan összefüggnek. Hogy a teljes kapcsolatrendszer kirajzolódjék előttünk, ahhoz már csak a terjedelem és a mélység kapcsolatát kell a fenti megállapításoknál részletesebben szemügyre vennünk.

A mondategységek mélységének mérsékelten erős kapcsolata a teljes szöveg és a tartalmi egység szószámával azt jelenti, hogy ebben a korpuszban a mondategységek közötti erőteljesebb alárendelő szerkesztésmód és a mondanivaló bővebb kifejtésének lehetősége együtt járt; vagyis aki nagyobb szövegterjedelem létrehozására volt képes, az képes lehetett arra is, hogy gondolatait ne csak tartalmi – logikai, hanem grammatikai eszközökkel is strukturálja. Valószínű azonban, hogy ez az összefüggés a fejlődőben levő nyelvhasználatra vonatkozik inkább, mint a felnőttekére, hiszen pl. nem vonatkozhat a sokat, de primitíven beszélők szövegeire. Ilyen adatközlő ebben a vizsgálatban is akadt, és ez nyilvánvalóan érezteti hatását a korrelációs adatban: nélkülük ez az összefüggés erősebb számszerűséggel jelentkezett volna.

A mondatrészek mélységének kapcsolatai közül a legmagasabb értékű a mondategységek szószámával mutatott mérsékelten erős összefüggés, amely szerint a mondategységen belüli szószám emelkedése együtt jár a mondatrészek mélyebb alászerkesztettségével. A mondategységre vonatkozó, a 2.2. pontban közölt alapadatok bizonyossága szerint a mondat belsejében levő szavak átlagos száma e korpuszban nagyjából állandó; túlnyomó többségük mondatszintű mondatrész, ezért az átlagos mélység értéke alacsony. Ha mégis kerülnek alárendelő szintagmák a mondatba, ezek determinánsaival valóban növekszik a mondat szószáma. E korrelációs érték viszonylagos alacsonysága azonban arra figyelmeztet, hogy e kapcsolatot az életkor, a műfaj és a téma mellett bizonyára befolyásolják az írásos és elhangzó mondat megszerkesztésének különbségei is.

A mondatrészek mélysége gyenge, szignifikáns kapcsolatban van a tartalmi egységek szószámával is, ez a kapcsolat azonban gyengébb, mint ami a tartalmi egységek szószáma és a mondategységek mélysége között mutatkozik. Ez azt jelenti, hogy a tartalmi egység bővebb kifejtése inkább a mondategységek, kevésbé a mondatrészek

közötti alárendelések számának növekedésével jár együtt, s ennek nyilvánvalóan ismét az élőnyelvi szerkesztésmód az oka.

Ha ezek után megfigyeljük, hogy a mondategység mélysége és a különböző szövegegységek közötti korrelációk ellentétes irányban erősödnek, mint a mondatrészek mélységével kialakult adatok, figyelemre érdemes tendenciát látunk. A mondategység mélységének alakulásában ugyanis a teljes szöveg adott idő alatt létrehozott terjedelme, a mondatrészekében pedig a mondategység szószáma a domináns tényező, vagyis: a legmagasabb szintű nyelvi egység (a szöveg) kifejtésének nagyobb mértéke nagyobb mélységű összetett mondatokkal jár együtt, a mondategység belsejének viszonyai pedig egymást befolyásolják. (A mondategységen belüli viszonyokat az 5. pontban, a nem szignifikáns adatok között újból tárgyalni fogjuk, megvizsgálendő, mi az oka annak, hogy e mutatók kapcsolatrendszere a leggyengébb.)

Végül tekintsük meg az összevont mélységmutató kapcsolatait is! Korrelációi erősödésének iránya a mondategységek ilyen adataival egyezik meg; legerősebb a kapcsolata a teljes szöveg, leggyengébb a mondategység szószámával. Ez az egyezés arra utal(hat), hogy az összevont mutató működésére a mondategységek mélysége erősebb hatást gyakorol, mint a mondatrészeké (bár ennek hatása is érződik benne). Ezt az összefüggést jól megvilágítja a két érintett mutató szórásértékének különbözősége: a mondategységek mélységének relatív szórása 8,06 %, a mondatrészeké 3,68 %, s ez azt jelenti, hogy ebben a korpuszban az adatközlők a mondategységek alárendelő viszonyainak megszerkesztésében lényegesen nagyobb (több, mint kétszeres) variációs lehetőséggel tudtak élni, mint az alárendelő szintagmák szerkesztésében (a kérdést részletesebben l. az alábbi pontban).

4. A mélységmutatók kapcsolatrendszere

E mutatók egymás közötti korrelációs adatai a következők:

- a mondategységek és a mondatrészek átlagos mélysége között: $-0,0412$,
- a mondategységek átlagos mélysége és az összevont mélységmutató között: $0,9076$,
- a mondatrészek átlagos mélysége és az összevont mélységmutató között: $0,3781$.

Az adatok közül az első igen kicsiny értékével azt jelzi, hogy a két jelenség nem függ össze egymással, a negatív előjel azonban arra enged következtetni, hogy egy csekély mértékű ellentétes tendencia mégis működik közöttük. A kérdés értelmezésére a nem szignifikáns kapcsolatok tárgyalásakor visszatérünk.

A mondategységek és a mondatrészek mélységének az összevont mutatóval kialakult korrelációs viszonya evidenciát fed, hiszen az összevont mutatót a másik kettő segítségével számoltuk ki. Így ez az adat összefüggésvizsgálatra nem használható; mégis

figyelemre érdemes, hogy az evidens esetekben várható 0,9 fölötti, igen erős korrelációs érték csak a mondategységek mélységével alakult ki, a mondatrészek mélységével mutatott értéke feltűnően alacsony. Ahhoz, hogy e jelenségre magyarázatot leljünk, érdemes áttekintenünk a mélységmutatók környezetét és alapstatisztikai adataikat egyaránt.

Ehhez először a három mélységmutatónak a teljes vizsgálatban kapott összes szignifikáns korrelációját foglaljuk táblázatba, benne azokat az adatokat is, amelyeket a NrNYt. XXXIII-ban megjelent írásban tárgyaltuk részletesen. (A korrelációs adatokat a jobb áttekinthetőség érdekében egy tizedesre kerekítettük.)

	mélység		
	<u>mondat</u> <u>egységek</u>	<u>mondat</u> <u>részek</u>	<u>összevont</u>
iteráció	0,5	0	0,4
szókincsgazdagság	0,3	0,3	0,4
elvont jeltíp. ar.	0,3	0,3	0,4
a teljes korpusz szövegszavai	0,5	0	0,5
a tart. egységek szövegszavai	0,4	0,3	0,5
a mondategység szövegszavai	0	0,3	0,2
a szemantikai hibák aránya	0	-0,3	-0,2
a kötőszók hibáinak aránya	0,2	0	0
az összes hiány aránya	0	-0,3	0

E táblázatból az derül ki, hogy az összevont mélységmutatóban nagyjából azonos súllyal van jelen a mondategységek és a mondatrészek átlagos mélysége, vagyis ezek az adatok nem magyarázzák a korrelációk fent taglalt nagy eltérését. Az alapstatisztikai adatok viszont közelebb visznek a lehetséges válaszhoz:

	mélység		%-os
	<u>mondategységek</u>	<u>mondatrészek</u>	<u>különbségek^{xx}</u>
átlag	1,33	1,17	12,06
első osztályközép	1,11	1,05	5,41
utolsó osztályközép	1,61	1,29	19,88
az osztályközök			
száma ^x	26	13	50,00
relatív szórás	8,06	3,68	54,34

^x az osztályközök mindkét adatsorban 0,02 értékkel emelkednek

^{xx} a %-os különbségek a mondategységek átlagos mélysége javára

Az alapadatok önmagukban is, a százalékos különbségek pedig azoknál sokkal szemléletesebben jelzik azt az eltérést, ami e két nyelvi szint mélységi viszonyait e korpuszban jellemzi (s amelyet a fentiekben már többször érintettünk): kétségtelennek látszik, hogy a 14 éves adatközlőink élőnyelvi szövegeiben a mondategységek alárendelő szerkesztésmódja erőteljesebb, mint a mondatrészeké. A jelenséget fentebb már magyaráztuk (a még nem teljesen fejlett nyelvhasználatra és az elhangzó beszéd mondat szerkesztési sajátosságaira hivatkozva); itt most ez a megfigyelés alkalmas arra, hogy indokoljuk vele a 4. pont elején közölt korrelációs adatok (0,9076 és 0,3781) közötti, elvileg érthetetlen különbséget: az összevont mélységmutató elsősorban a (mondatrészekhez képest) erőteljesebb alárendelő szerkesztésmódú mondategységek mélységi viszonyait tükrözi.

A továbbiakban – a teljesség kedvéért – megismételjük a mélységmutatók fentebb tárgyalt főbb szignifikáns összefüggéseit:

- a mondategységek alárendeltségi fokának emelkedése együtt jár az iterációs tendencia csökkenésével;

- a mondatrészek mélysége a szókészlet nagyobb választékával mutat összefüggést;

- az elvont jeltípusok arányának szövegbeli növekedése maga után vonja mind a mondategységek, mind a mondatrészek erőteljesebb alárendelő szerkesztésmódját;

- az adott idő alatt létrehozott szöveg terjedelme a mondategységek mélységi viszonyaival függ össze;

- a tartalmi egységek terjedelme a mondategységek mélységével erősebben, a mondatrészekével gyengébben hozható kapcsolatba;

- a mondategységen belüli összefüggéseket jól jellemzi a mondategység szószáma és a mondatrészek mélysége közötti összefüggés.

5. A vizsgált mutatók közötti nem szignifikáns viszonyok a következők:

1. iteráció – mondatrészek mélysége: 0,0013
2. az elvont jeltípusok aránya – a mondategység szószáma: 0,1025
3. a teljes szöveg terjedelme – a mondatrészek mélysége: 0,1571
4. A mondategységek mélysége – a mondategység szószáma: 0,0975
5. a mondategység mélysége – a mondatrészek mélysége: -0,0412

Az adatokból rendkívül tiszta kép bontakozik ki: a mondategység belsejét vizsgáló két mutató (a szószám és a mondatrészek mélysége) nincs összefüggésben az iterációval, az elvont jeltípusok arányával, az adott idő alatt létrehozott szöveg terjedelmével és a mondategységek mélységével.

Tekintsük meg ezeket az adatokat részletesebben is!

A mondategység szószámának alakulása a 2. adatsor szerint nincs kapcsolatban a szöveg elvont jelentésű jeltípusainak arányával, bár egy csekély mértékű egyenes arányosságot mégis kiolvashatunk belőle. Ez – mint fentebb már kifejtettük – alátámasztja azt a megfigyelésünket, hogy a mondategység szószáma adott életkorban és közlési körülmények között (valószínűleg főleg az elhangzó beszédben) viszonylag állandó, ám a legmagasabb absztrakciós szintű szókincs birtokosai az átlagosnál több szóból képesek mondategységeiket felépíteni.

Ennél érdekesebb a mondategységek mélysége és terjedelme közötti viszony (4. adatsor), amely azt jelzi, hogy a mondategységek közötti alárendelő struktúrák gyakorisága nincs összefüggésben a bennük lévő szavak számával (bár egy igen gyenge egyenes arányosság fennáll közöttük). Ez a megfigyelés ellentétes B. Fejes Katalinével (1987), aki általános iskolás tanulók írásos szövegeinek statisztikai elemzése alapján arra a következtetésre jut: „Ha hierarchikusabban tagoltak a szöveg mondategységei..., akkor a főmondatok mondatrészei zsúfoltabbak”, vagyis: ha emelkedik a mondategységek közötti mélység, emelkedik a főmondatokban levő szavak száma is (B. Fejes 1987, 34). Ha ez a két, statisztikai úton nyert megfigyelés egymással szembekerül, mindkettő igazságának feltételezése mellett meg kell találnunk az eltérés okát. Véleményünk szerint ez az ok a kétféle közlésfajta mondat szerkesztési különbségeiben keresendő: az írásbeli mondat megszerkesztésében sokkal nagyobb tere lehet a balra bővülő szerkezetes mondatrészeknek anélkül, hogy ezek a megértést akadályoznák, mint az élőszóban elhangzó mondat esetében; így ez utóbbi természetszerűleg kevesebb szóból tevődik össze még akkor is, ha a spontán élőnyelvben oly gyakori témaismétlő szerkezetek, halmozott mondatrészek és módosító mondatrészetek fő- és mellékmondatban is lehetséges felbukkanása még emeli is a szószámot. Így azt mondhatjuk, hogy a mondategységen belüli szószám és a mondategységek közötti alárendelések száma valóban egyenes arányosságban áll, de ez a tendencia az élőnyelv sajátos szerkesztésmódja miatt nem érvényesül meghatározó mértékben, csak a legjobb mondataiban. Az

élőnyelv fent említett, a mondategység szószámát emelő jellemzőről l. Bicskeiné 1985. 61–2. (Meg kell azonban jegyeznünk, hogy a B. Fejes által talált összefüggés a főmondatok szószámára vonatkozik, az itt közölt adatok pedig általában a mondategységekre.)

A mondatrészek mélysége nincs összefüggésben az iterációval, a szöveg terjedelmével és a mondategységek mélységével. A nullához közelítő korrelációs adat lényegében minden összefüggést kizár a szavak ismétlési/ismétlődési tendenciája és a mondatrészek alárendeltségi foka között (l. adatsor). Ez akkor válik könnyen elfogadhatóvá, ha szemügyre vesszük az iteráció korrelációt a különböző nyelvi szinteken:

a teljes szöveg szószámával:	0,7641
a tartalmi egység szószámával:	0,4843
az összetett mondaton belüli viszonyokkal:	0,4565
a mondategység szószámával:	0,2718
a mondatrészek közötti viszonyokkal:	0,0013

Így az adatok bizonyítják az empirikus megfigyelést: a szóismétlési jelenség a szöveg makroszintjén és a tartalmi egységek szintjén a legerősebb, de már a bekezdés sem viselheti el olyan arányban, mint a teljes szöveg; a tartalmi egységek szószáma és az egységet felépítő mondategységek mélységi viszonyai felől nézve közel azonos erősségű, hiszen e két tényező azonos nyelvi szint két különböző aspektusa; ennél lényegesen gyengébb e tendencia a mondategységet alkotó szavak körében, hiszen valóban ritka az, amikor egy mondategységen belül fogalomszó ismétlésére/ismétlődésére kerül sor, ám ott is van némi valószínűsége a névelők, névutók, határozószók és módosító elemek ismétlődésének – ezért a gyenge korrelációs adat itt indokolt. E sor végén jutunk el a szerkezetes mondatrészekhez, és zárhatjuk ki annak a valószínűségét, hogy ezekben szóismétlés lehetne (ismétlődés is aligha lehet más, mint a névelő).

A mondatrészek között megjelenő alárendelő viszonyok aránya nem függ össze az adott idő alatt létrehozott szöveg terjedelmével sem (3. adatsor). Ez jelzi: egyáltalán nem biztos, hogy a bőbeszédű, közlékeny ember mondataiban több lenne a szerkezetes mondatrészek száma, bár a szignifikancia alsó határához közelítő korrelációs adat halványan sejtet egy ilyen tendenciát.

Nincs szignifikáns kapcsolat a *mondatrészek és a mondategységek mélységmutatója* között sem (5. adatsor), sőt a korreláció értéke oly kicsiny, hogy szinte minden kölcsönösséget kizár; fel kell azonban figyelniük az előjelre, mivel azt mutatja, hogy a két nyelvi szinten levő alárendelő szerkesztésmód ellentétes tendenciájú. Valószínűnek tartjuk, hogy ez az ellentéteség teljesen hiányzik a művelt felnött írásos közleményeiből, a fejlődő nyelvhasználat spontán elhangzó szövegeit azonban annál inkább jellemzi; azaz az írott szöveg mondatrészeinek mélysége a beszélt nyelvben inkább a mondategységek mélységében jelenik meg. Saját elemzési tapasztalatainkból ennek az ellentétes tendenciának bizonyítékát látjuk a körülírások szerkesztésmódjában

(is). Ez az élőnyelvben, ott is főleg a fejletlen szókincsű szövegekben előforduló konstrukció ugyanis pontosan a fent vázolt módon épül fel: egy-egy, a mondatból szókincs-hiány (vagy pillanatnyi zavar) miatt kimaradó szót úgy próbál pótolni, hogy annak fogalmi jegyeit különálló, gyakran alárendelt mondatokba foglalja (BICSKEINÉ 1984. 19–20.).

A dolgozatban közölt adatokhoz a már publikált részekkel együtt szorosan hozzátartozik még az itt ismertetett mutatóknak az élőnyelvi töltelék- és hibajelenségekhez fűződő kapcsolatrendszere; ám ez már egy másik dolgozat témája. E vizsgálat (részben vagy egészében bizonyára vitatható) mutatói korrelációs adataikkal (akár szignifikánsak, akár nem azok) feltártak egy sor nyelvi összefüggést, amelyek között voltak nyilvánvalók is, meg olyanok is, amelyek figyelmet érdemelnek. A kutatás elvégeztével azonban bizonyítottnak látjuk, hogy ilyen – lényegében egyszerű, vagy a szakirodalom nyomán haladva, de leegyszerűsített – mutatókkal jól lehet jellemezni a tanulók spontán beszédtevékenységét, és hogy bizvást számíthatunk az ilyen jellemzés során arra, hogy feltárulnak az egyéni nyelvhasználatnak azon elemei, amelyek igénylik az anyanyelvi képességfejlesztést.

IRODALOM

- B. Fejes Katalin 1987, A szintmélységmutató környezetstruktúrája. *Acta Academiae Paedagogicae Szegediensis, Series Linguistica, Litteraria et Aestetica*. Szeged, 29–35.
- Bicskei Dezsőné 1984, A mondat hosszúság és a mondategészbeli kapcsolások 14–15 éves gimnazisták nyelvhasználatában (élőszóbeli és írásos szövegek alapján). *Néprajz és Nyelvtudomány* 28, 17–28.
- Bicskei Dezsőné 1985, Középiskolások élőbeszédének mondatszerkezeti jellemzői. *Magyar Nyelvőr* 109, 48–62.
- Bicskeiné Zsulán Julianna 1990, A szóbeli kifejezőképesség minőségének vizsgálata a szókincs elemzés segítségével. *Magyar Nyelvőr* 114, 67–85.
- Deme László 1971, *Mondatszerkezeti sajátosságok gyakorisági vizsgálata*. Budapest, Akadémiai.
- Károly Sándor 1970, *Általános és magyar jelentéstan*. Budapest, Akadémiai.
- Kovacsics József 1979, *Statisztika*. Budapest, Tankönyvkiadó.
- Nagy Ferenc 1980, *Kriminallisztikai szövegnyelvészet*. Budapest, Akadémiai.
- Wacha Imre 1988, Élő nyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz (A gazdagréti kábeltelevízió élő nyelvi felvételei)

alapján). In: Kontra Miklós szerk., *Beszélt nyelvi tanulmányok Linguistica Series A/1*. Budapest, MTA Nyelvtudományi Intézete, 102–158.

Zsilka Tibor 1974, *Stilisztikai és statisztika*. Budapest, Akadémiai.

Zsulán Julianna 1989–90, Az élőnyelvi szöveg hiba- és töltelékjelenségeinek belső összefüggései. *Néprajz és Nyelvtudomány* 33, 91–103.

**ON SOME INTERNAL RELATIONS IN ORAL DISCOURSES
OF FOURTEEN-YEAR-OLD CHILDREN**

JULIANNA ZSULÁN

The paper presents the results of a longer project on spontaneous oral discourses of fourteen-year-old children. A statistical analysis is used to determine the relations between the data concerning qualitative and quantitative composition of the corpora, the sizes of text units at different levels, and the subordination of clauses and phrases. The goal of this analysis is twofold: to examine the characteristics of the dynamism of indices as well as their appropriateness in giving an objective measure to developing language use.

The most important observations are the following: (1) in texts with fewer-than-average words per phrase word repetitions tend to be more frequent; (2) word repetitions are independent of phrasal hierarchy; (3) the role of expressions of higher abstraction is more prevalent in the explication of paragraph size texts than in whole texts; (4) the extent and quality of individual vocabulary and the structures embedding clauses and phrases are interdependent; (5) the size of clauses is related to the depth of phrases; (6) the depth of phrases is not related to the size of text or the depth of clauses or iteration.