

P. TÓTH TEODÓRA

ELTE Eötvös Loránd Tudományegyetem, Nyelvtudományi Doktori Iskola
tothariannateodora@kancellaria.elte.hu
<https://orcid.org/0000-0003-2686-3448>

P. Tóth Teodóra: Oksági viszonyt jelölő kötőszavak korpuszalapú vizsgálata 14–15 éves diákok érvelő fogalmazásaiban

A corpus-based study of causal conjunctions in argumentative essays written by 14–15-year-old students

Alkalmazott Nyelvtudomány, Különszám, 2025/1. szám, 237–255.

doi: <http://dx.doi.org/10.18460/ANY.K.2025.1.013>

Oksági viszonyt jelölő kötőszavak korpuszalapú vizsgálata 14–15 éves diákok érvelő fogalmazásaiban

A corpus-based study of causal conjunctions in argumentative essays written by 14–15-year-old students ANY_elte_0000-0003-2686-3448_tothariannateodora@kancellaria.elte.hu

The last problem of the Hungarian section of the central written entrance exams for eighth-form pupils applying to secondary schools has been the composition of an argumentative text every year since 2013. In argumentative texts, there is a need to use conjunctions marking logical relationships (including ones marking causal relationships) in an accurate and nuanced way (Adamik et al., 2005: 517). I classified the causal relationships able to be marked by a conjunction in compound and complex sentences into four categories (cause-effect, effect-cause, conditional subordinate clause, inferential subordinate clause) following the categorisation in Magyar grammatika (Keszler, 2000: 273).

The goal of the study is to observe the distribution of conjunctions marking causal relationships in argumentative texts of 14–15-year-old pupils, and to compare that to data gathered from a representative corpus. The main research questions are:

1. What is the distribution of the conjunctions marking causal relationships in the texts written by the pupils?
2. In what proportion do the conjunctions fall into the four categories of causal relationships?
3. What similarities and differences can be observed between the distributions of conjunctions in pupils' essays and the representative corpus chosen as a reference corpus?

The study is comprised of two phases: the first phase consists of the analysis of pupils' essays, while the second consists of the study on the representative corpus as well as the comparison of the results.

In the first phase of the study the sample consists of 30 texts. These were chosen from the material of one of the author's earlier studies (P. Tóth, being published). The pupils participating in the study completed an exercise set conforming to the requirements of the central written entrance exams, carrying its typical exercise types and structure. The 30 documents chosen were collected from pupils of the same Budapest primary school.

In the first phase, the conjunctions found in the essays were listed using the e-magyar part-of-speech tagging software, and after checking and correcting the annotations, the conjunctions marking causal relationships were selected. The frequency of these was noted in the whole corpus (counting all appearances as well as the ones marking causal relationships). In the second phase, the spoken language, official, press, personal-forum and personal-social subcorpora (MNSZ-H) together as well as the scientific subcorpus (MNSZ-T) were analysed separately from the Magyar Nemzeti Szövegtár (MNSZ) which was chosen as the reference corpus.

Out of the 430 conjunctions found in the pupils' texts, 123 mark causal relationships. *Ha* and *hogyha* (that start conditional subordinate clauses) constitute almost half of all cases (53), followed by the type connecting dependent or independent clauses in an effect-cause order (38 cases). The cause-effect order (implying deductive thinking) appeared in 16 cases, and finally, conjunctions starting inferential

subordinate clauses appeared in 14 cases in the texts. In the reference corpus, *ha*, *hogy* and *hogyha* appeared with lower, while other conjunctions appeared with higher relative frequencies compared to the essays. The essays are less varied in their usage of conjunctions compared to the MNSZ subcorpora. The analysis of the 30 essays showed that in pupils' written texts, conjunctions beginning conditional subordinate clauses constitute an exceptionally high proportion of conjunctions marking causal relationships (especially *ha*), while the proportion of conjunctions connecting independent clauses in a cause-effect order (which imply deductive reasoning) is very low, although there were several different conjunctions of that function in the texts.

The usage of conjunctions by the pupils is observed to be vastly different from the distributions gathered from the MNSZ subcorpora, but there would be a need to repeat the study on a bigger sample (preferably with several essays from the same child, possibly of various text types) and more annotators to more thoroughly support this observation and discover the causes thereof. The results of the analysis of conjunctions of causal relationships in written texts and the categorisation system developed can be applied in the areas of subject didactics and pedagogical assessment, applied rhetoric as well as language technology development.

Keywords: written composition, argumentative text, corpus linguistics, conjunction, causal relationships

1. Bevezetés

1.1. A fogalmazásfeladatok értékelésének problematikája

A szövegalkotási feladatok értékelése (beleértve az iskolai környezetben végzett fejlesztő és diagnosztikus értékelést, illetve a szövegalkotási képességre vonatkozó méréseket) mind tantárgy-pedagógiai, mind kutatómódszertani szempontból számos problémát vet fel. Ennek egyik oka a szövegalkotási képesség komplexitásában keresendő, emellett azonban már a feladat kijelölésekor vagy a mérés tervezésekor, módjának megválasztásakor, illetve az eredmények megbízhatóságának értékelésekor is figyelembe kell venni a szövegalkotás folyamatát befolyásoló egyéb tényezőket: a kiadott feladat sajátosságait, a mérés körülményeit a tanulók pillanatnyi állapotát (Csákerényiné Tóth, 2011). A szövegalkotási képességre vonatkozó átfogóbb felmérésnek az is akadály, hogy „egyes tanulók különböző típusú, műfajú szövegekre vonatkozó fogalmazási feladatokban eltérő színvonalon teljesítenek” (Molnár, 2000: 51–52).

Más tudásszintmérő tesztekkel ellentétben a fogalmazások értékelésében még részletesen kidolgozott értékelési skálák, útmutatók alkalmazása mellett is nagyobb szerepet játszik az szubjektivitás, a bíráló értelmezése, szigorúsága (Nagy, 2013). Juhász Milán (2021) érvelő esszék értékelésének szempontjaival kapcsolatos vizsgálata azt is kimutatta, hogy az eltérő tanítási tapasztalattal rendelkező pedagógusok különböző jelentőséget tulajdonítanak az egyes kritériumoknak. A szövegalkotási képesség színvonalát a szövegalkotás produktumai, a fogalmazások alapján lehet megítélni. A mérés hitelessége fokozható, ha egyes tanulóknak több írásművét vonjuk be a vizsgálatba, illetve, ha az egyes fogalmazásokat több bíráló értékeli.

A fogalmazások értékelésének iskolai gyakorlatában leggyakrabban tartalmi és nyelvi szempontok szerinti osztályzatot kapnak a tanulók (Molnár, 2000: 49), míg a magyar nyelv és irodalom érettségén hármass felosztás (tartalom, szerkezet,

nyelvi megformáltság) szerint történik az értékelés. Magyarországon nincsen rendszeres országos szövegalkotási mérés, azonban számos kutatás foglalkozott különböző korosztályok szövegalkotási képességének, illetve írott szövegeinek vizsgálatával. Ezek közül kiemelendők Orosz Sándor (1972), Kádárné Fülöp Judit (1990) és Horváth Zsuzsanna (1998) reprezentatív mintán végzett kutatásai, amelyek részletes értékelési szempontrendszeret alkalmaztak. Orosz Sándor (1972) a magyar szakirodalomban elsőként olyan szempontrendszer kialakítására tett kísérletet, amely minél objektívabb értékelést tesz lehetővé. Ennek érdekében három csoportba sorolva (tartalmi, szerkesztési és stilisztikai minőség) tizennyolc kritériumot határozott meg. Az IEA (International Association for the Evaluational Achievement, Nemzetközi Iskolai Eredménykövető Társaság) által felkért nemzetközi szakértői munkacsoportja (lásd Kádárné Fülöp, 1990: 9) által kidolgozott és Magyarországon is elvégzett mérésének keretében alkalmazott standardizált értékelési rendszerét az összbenyomás mellett hat analitikus értékelési szempont alkotja: tartalom, szerkesztés, stílus, nyelvhelyesség, helyesírás, külalak. Ennek a mérésnek különös érdekessége, hogy az említetteket követően tartalmaz egy további szempontot: az értékelő személyes reakcióját, amely nem a tanulói teljesítmény mutatójaként, hanem „a bíráló érzelmeinek levezető csatornájaként” szolgált. (Kádárné Fülöp, 1990: 47–48) Ezt a szempontrendszert (és az IEA-felmérés mérőeszközeinek átdolgozott változatát) használta Nagy Zsuzsanna 2009-ben végzett vizsgálatában (Nagy, 2013), de a hat analitikus értékelési szempont feltűnő hasonlóságot mutat a középiskolai írásbeli felvételi ötelemű pontozási rendszerével is (lásd 1.2. fejezet). Horváth Zsuzsanna 1998-as kutatásában harminckét darab, négyfokozatú skálán értékelt itemből álló szempontsorral dolgozott, amelyek az összbenyomáson kívül tartalmi, szerkezeti és nyelvi aspektusokra terjedtek ki. A szempontok között szerepelt többek között az érvek száma, tartalma, a címadás és a saját állásfoglalás. Ennek a kutatásnak jelen munka szempontjából legfontosabb tanulsága az értékelés objektivitására vonatkozik: az ismeretlen tanulók fogalmazásait a pedagógusok objektívabban értékelték, mint az ismerős diákok munkáit.

1.2. Érvelő fogalmazás a középiskolai felvételin

Magyarországon 2004-ben hirdettek először egységes írásbeli felvételi vizsgát a középiskolába jelentkező tanulók számára. Az 50 pontos anyanyelvi vizsgarész nyelvhasználati, nyelvtani és szövegértés, valamint szövegalkotási feladatból áll. Ez utóbbi „összetettebb tizedik feladat önálló szövegalkotást kíván a tanulóktól: adott témáról, adott műfajban kerek, egész, lezárt szöveget kell fogalmazniuk. A fogalmazásnak a téma megtartása és a terjedelmi kíváncsalom mellett (ez évfolyamonként eltérő) az alapvető nyelvhelyességi, szerkesztési, helyesírási és stílusbeli követelményeknek is meg kell felelnie” (Tájékoztató 2023).

Ez, a 2010-es évek közepétől tipikusnak tekinthető felvételi fogalmazás egy 10–12 mondatos, tartalmilag és formailag szerkesztett, összefüggő, lezárt szöveg,

amely jellemzően két felvetett álláspontot alátámasztó 2-2 érvet tartalmaz, ritkábban egy állítással kapcsolatos két érvet és két ellenérvet, emellett fontos tartalmi eleme a tanuló saját álláspontjának megfogalmazása. Egy ilyen feladat látható a 2. ábrán:

1. ábra: A 2022. január 22-i 8. évfolyamos központi írásbeli feladatsor feladata (OH 2022: 9)

Sokszor felmerül a kérdés, hogy szükség van-e manapság a felső tagozaton a nyomtatott tankönyvekre. 10–12 mondatos fogalmazásodban fejtsd ki erről a **véleményed**et, mutass be **2-2 érvet** a nyomtatott tankönyv használata **mellett és ellen!** A **saját álláspontodat is fogalmazd meg**: Neked **szükséged van** a nyomtatott tankönyveidre?

Figyelem! A 4 mondatnál rövidebb fogalmazás nem értékelhető! Ha a fogalmazás nem éri el a 10 mondatot, az jelentős pontlevonással jár.

Az évről érve megjelenő hivatalos javítási útmutató a következő értékelési szempontokat határozza meg: tartalom (3 pont); a szöveg szerkezete (3 pont); stílus, szóhasználat, nyelvhelyesség (1 pont), helyesírás (2 pont), külalak, íráskép (1 pont). Bár a fogalmazások terjedelme 2007 óta egységesen 10–12 mondat, csak 2018 óta határoz meg szigorú terjedelmi minimumot a javítókulcs: 4 mondatnál rövidebb dolgozat esetén egységesen minden itemre 0 pontot kap a felvételiző, 10 mondatos terjedelem alatt pedig az összpontszám legfeljebb fele adható. Az utóbbi esetben a pontlevonás a szerkezet, a stílus, szóhasználat, nyelvhelyesség és a helyesírás itemekből történik, a tartalmi pontszámot nem érinti.

A stílus, szóhasználat item pontozása tehát függ a fogalmazás terjedelmétől, mivel azonban ezt mondatzámban határozza meg mind a feladat szövege, mind a javítókulcs, a javítók számára tapasztalatom szerint rendszeresen komoly dilemmát okoz a 8–9 mondatos (tehát az elvárt terjedelemtől rövidebb), ugyanakkor igényes nyelvezetű, jól sikerült fogalmazások értékelése. Az összetettebb mondat szerkezetek miatt ezek a dolgozatok gyakran a tagmondatok vagy szavak számát tekintve hosszabbak, mint számos 10 mondatot elérő, azonban zömmel egyszerű vagy egyszerűen összetett mondatokból álló szöveg. A terjedelem mellett a javítókulcs három kritériumot szab meg a stílus, szóhasználat item értékeléséhez: az erre járó 1 pont abban az esetben adható, ha a dolgozat stílusa, szóhasználat:

- megfelel a témának, a köznyelvi normának;
- gazdag szókincsű;
- nyelvhelyességi szempontból kifogástalan (Javítókulcs 2024).

A fenti három szempont mindegyike számos kérdést vet fel: Hogyan definiálható a köznyelvi norma; hogyan viszonyul egymáshoz, mennyire élesen válik el az írott és a beszélt nyelvi norma; reális elvárás-e a 14–15 éves diákokkal

szemben az ezek közötti különbségtétel, kódváltás (különösen fontos kérdés ez a nyelvjárási háttérrel rendelkező tanulókkal kapcsolatban)? Hogyan befolyásolhatja az értékelést a tanuló és a tanár közötti kisebb vagy nagyobb életkori különbség, a nyelvhasználatban is megmutatkozó generációs szakadék? Mi tekinthető gazdag szókincsnek; miben és hogyan mérhető le a szókincs gazdagsága (a felvételi javításakor rendelkezésre álló időkeret és eszközök figyelembevételével)? Mit jelent a kifogástalan nyelvhelyesség; beszélhetünk-e a magyar nyelv valamennyi területi változatára érvényes nyelvhelyességi szabályokról?

1.3. Fogalmazáskészség fejlődése

Az elvárások, értékelési szempontok mérlegeléséhez elengedhetetlen az írásbeli szövegalkotási képesség fejlődési folyamatának rövid áttekintése, hiszen ez a tanulási folyamat nem csak a 8. évfolyamos felvételiző tanulók esetében, de valójában még a felsőoktatásba belépő fiatal felnőttek esetében sem tekinthető lezártnak. Az írásbeli szövegalkotási képesség fejlődésének ütemét az iskolai fejlesztés, valamint az ennek keretében alkalmazott írásgyakorlatok jellege és felépítése mellett erősen befolyásolja a nyelvi tudás szintje is, amelyben már az iskolakezdekéskor nagy különbségek mutatkoznak az egyes tanulók között. „Idővel valamennyi gyermek írásbeli szövegalkotási képessége fejlődik, de nem egyforma ütemben és nem azonos módon. (...) Ugyanakkor nagyon sokan eljutnak arra a szintre, hogy – ha nem is kifogástalan módon, de – írásba tudják foglalni közlendőjüket, és némi rutinra tesznek szert ebben” (Tóth, 2006: 6).

Bereiter és Scardamalia (1987) modellje szerint az írásstratégia szempontjából a kezdő írókat a tudáselmondó (knowledge telling), míg a gyakorlott írókat a tudás-átalakító (knowledge transforming) stratégia jellemzi. Míg az előbbi a 7–10 éves korosztályra jellemző (ezt megelőzően az első próbálkozások, az előkészítés időszakáról beszélhetünk, amelyben még a gondolatok írásban való rögzítése jelenti a legfőbb kihívást), az utóbbi 10 és 14 éves kor között épül ki. A tudáselmondás stratégiájára a szóbeliségből ismert, kész elbeszélési minták alkalmazása, a sémák és forgatókönyvek megjelenése jellemző. A tudás-átalakítás szakaszában a tartalom és az olvasó felé fordul a figyelem, a műfaji elvárások mellett megjelennek a retorikai szituációból adódó feltételek is a szövegalkotás folyamatát befolyásoló tényezők között; következetesebbé, pontosabbá válik a szövegkohéziós eszközök és az igeidők használata (Becker-Mrotzek & Böttcher, 2014: 54–63; magyarul lásd Major 2015: 231). „Az érveléstechnika is fejlődik, megjelenik az időrend mellett a tematikus csomópontok logikája szerint történő szervezés és a pro-kontra-séma” (Major, 2015: 231). Bár az egyes írásstratégiák közötti váltás nem elsősorban az életkor függvényében, hanem a tapasztalat, a metakognitív tudás meglététől függően megy végbe, az egységes tantervek szerint szerveződő iskolai oktatás feltételezi és elvárja, hogy a tanulók többsége hasonló életkorban érje el a fejlődés egyes szakaszait, tudjon megbirkózni az egyre

összetettebb szövegalkotási feladatokkal. A fentiekkel összefüggésben a középiskola felvételin a 4. és 6. évfolyam számára kijelölt fogalmazási feladat jellemzően elbeszélés írása, amely nagyobb mértékben támaszkodik a beszélt nyelvből, illetve a szépirodalomból ismert forgatókönyvek alkalmazására és a kronologikus szerkesztésre, míg a 8. évfolyam esetében megszokott érvelő szöveg alkotásához „a megfelelő fogalmazáskészségen túl elengedhetetlenül szükséges egyfajta elemző, rendszerező szemlélet, a kritikus gondolkodás, a logikai következtetés, valamint az objektív és a szubjektív nézőpont közötti különbségtétel képessége, az érvek feltalálásában és nyelvi megformálásában való jártasság, a megfelelő nyelvi sémák ismerete és alkalmazása” (P. Tóth, megjelenés alatt). Ez a szövegtípus ugyanis „(...) objektív és szubjektív elemek összekapcsolását várja el: az összegyűjtött információk megfogalmazásának, „tálalásának” egyéni módját, előzetes tudáson, tájékozottságon alapuló saját gondolatokat. (...) A diák szabatosan megfogalmazva ne csupán minősítsen, meséljen, hanem a véleményét hiteles, tudományos érvekkel is támassza alá” (Tolnai Varga, 2018: 29). Fontos azonban megemlíteni, hogy noha a tanulmány írásakor érvényben lévő Nemzeti alaptantervben (NAT 2020) a célok között szerepel, hogy a felső tagozaton a tanulási folyamat tervezésekor szempont kell legyen a középiskolai felvételi követelményrendszer, a 8. évfolyamig bezárólag kevés szó esik a dokumentumban az érvelésről, az alsó és a felső tagozat végére elérendő mérőföldkövek között alig látunk különbséget. Nem jelenik meg konkrét célként az érvelő típusú fogalmazás írása a szövegalkotáson belül, illetve fontos változás, hogy eltűnik egy, a 2012-es NAT-ban még szereplő elvárás: a vélemény és az igazságok, tények közötti különbségtétel képessége; a kritikus gondolkodás fogalmát pedig a mérlegelő gondolkodás váltja fel.

A szövegtípust érintő korábbi kutatásom rámutatott arra, hogy míg a tanulók többségének sikerült megfelelő számú érvet összegyűjtenie, az érvelés részműveletei közül különösen logikai kapcsolatok jelölése szempontjából gyengébbek a diákok dolgozatai (P. Tóth, megjelenés alatt).

1.4. Logikai kapcsolatok a szövegben: Az ok-okozati viszony jelölésének lehetőségei

Az előző alfejezetben említett írásstratégiai modell szerint a jellemzően 10–14 éves kor között kiépülő tudás-átalakító stratégiát alkalmazó íróra többek között a szövegkohéziós eszközök tervszerűbb, következetesebb használata jellemző. Az érvelő szövegtípusban kiemelt szerepet játszik az ok-okozati viszonyok megfelelő jelölése, hiszen ezek teszik az olvasó számára világosan követhetővé a gondolatmenetet (SRDP 2018), már Arisztotelész is tárgyalja a Rétorika III. könyvében is a kötőszavak helyes alkalmazásának jelentőségét (Arisztotelész, 1982: 1407b). A formális logikában (a természetes nyelv többértelműségét kizárandó) ezt a szerepet operátorok, matematikai szimbólumokkal is jelölhető, úgynevezett logikai kötőszavak töltik be (Bánki, 2004), az érvelő szövegalkotás

azonban sokkal inkább a retorikához, illetve az informális logikához áll közel: A formális logika szerint a következtetések érvényességét kizárólag a logikai sémák határozzák meg, míg a Hamblin és Walton által képviselt informális logika a formális logika eredményinek felhasználásával, de tágabb keretben, természetes nyelven megfogalmazott érvelésekkel foglalkozik (Margitay, 2007: 18), így a természetes nyelv kifejezéseinek jelentését, tartalmát, valamint az emberi tapasztalat jelenségeit is figyelembe veszi. (Adamik et al., 2005: 338)

A folytatásos szövegmondatok közötti tartalmi-logikai kapcsolat sok esetben jelöletlen: éppen az oksági viszonyt kifejező következtető és magyarázó mellérendelésre jellemző, hogy „elkötőszótlanodó” (Deme, 1965: 299), habár grammatikai szempontból a kötőszó elhagyása ezekben az esetekben hiányként jelentkezik (MG, 2000: 535). A kapcsolatok explicit jelölése vagy éppen nyelvi jelöletlensége a szövegalkotóra jellemző stilisztikai vonás lehet (Szikszainé Nagy, 2006: 195), azonban éppen az érvelő szövegekben szükség van a logikai viszonyokat jelölő kötőszók pontos, árnyalt használatára (Adamik et al., 2005: 517).

A Magyar grammatika (2000: 273) a kötőszók által jelöltté tehető alá- és mellérendelő szerkezetekben megjelenő oksági viszohnak négy altípusát különíti el, a csoportosítást az 1. táblázat mutatja be kiegészítve a következtető és az okadó magyarázó mellérendelő összetett mondatok további jellemző kötőszóival (MG, 2000: 545)

1. táblázat: Az ok-okozati viszonyok különböző kötőszókkal jelölhető altípusai

Típus funkció szerint	Jellemző mondat szerkezet	Jellemző kötőszók
Ok-okozati sorrendben kötik össze a tagokat	következtető mellérendelés,	<i>úgyhogy, tehát, így, ezért, ennél fogva, következésképpen, ennek következtében</i>
Okozat-ok sorrendben kötik össze a tagokat	okhatározói alárendelés, (okadó) magyarázó mellérendelés	<i>mert, merthogy, mivel, mivelhogy, minthogy, ugyanis, tudniillik, hisz, hiszen, úgyis, úgyse, másképp, különben, máskülönben</i>
A feltételes sajátos jelentéstartalom jelölője, amennyiben az oksági viszony feltételes tartalommal együtt jelenik meg	feltételes mellékmondat	<i>ha, hogyha, amennyiben</i>
Szándéktalan következményt jelöl az <i>úgy, olyan, akkora, annyi</i> stb. utalószóval társulva	következtető mellékmondat	<i>hogy, míg</i>

A kötőszavak mellett más szófaji besorolású szavak, valamint több szóalakból álló kifejezések is jelölhetik a szövegbeli logikai viszonyokat akár tagmondatok, akár a szöveg nagyobb egységei között. Ezek korpuszalapú vizsgálata részben alaktani sokszínűségükből fakadóan, részben abból adódóan vet fel számos problémát, hogy ezek alkalmazása nagyobb teret enged az egyéni nyelvi leleménynek, várhatóan egy ilyen vizsgálat sokkal részletesebb annotációs protokollt, valamint a nagyobb bizonytalanságból fakadóan mindenképp több annotátor bevonását igényelné.

1. 5. A kutatás célja és kérdései

A kutatás célja egyrészt az ok-okozati viszonyt jelölő kötőszavak eloszlásának vizsgálata 14–15 éves diákok érvelő fogalmazásaiban, valamint ezek összehasonlítása egy reprezentatív korpuszból nyert adatokkal; másrészt módszertani kísérletként az argumentatív szövegek nyelvi jellemzőire vonatkozó annotációs lehetőségek feltérképezése, egy lehetséges annotációs eljárás tesztelése.

A kutatás fő kérdései:

- (1) Milyen gyakorisággal fordulnak elő a kutatás anyagát alkotó 14–15 éves diákok által írt érvelő dolgozatban az egyes oksági viszonyt jelölő kötőszók?
- (2) Milyen arányban fordulnak elő az oksági viszony négy alkategóriájához köthető kötőszók a kutatás korpuszát alkotó, 14–15 éves tanulók által írt érvelő fogalmazásokban?
- (3) Milyen hasonlóságok és eltérések figyelhetők meg a kötőszók eloszlásában a 14–15 éves diákok esszéiben, illetve a referenciakorpuszként választott reprezentatív korpuszban?

2. A kutatás módszertana

2. 1. A kutatás anyaga

A kutatás két fő fázisra tagolódik: az első szakaszt a diákesszé elemzése, a másodikat a reprezentatív korpuszon elvégzett vizsgálat, illetve az eredmények összevetése alkotja.

A vizsgálat első fázisában 30 darab fogalmazás alkotja a mintát. Ezeket korábban elvégzett kutatásom anyagából válogattam (P. Tóth, megjelenés alatt). A kutatásban részt vevő nyolcadikos diákok a központi írásbeli vizsga követelményeinek, szokásos feladattípusainak, felépítésének megfelelő feladatsort töltöttek ki a felvételi szokásos kereteihez hasonló helyzetben. (A tanulók 45 percig dolgozhattak önállóan, csak a felvételi vizsgán is megengedett segédeszközök használatával, viszont ismerős környezetben, a saját iskolájukban.) A dolgozatokat a diákok 2019. október és 2020. január között írták meg, tehát az adatfelvétel a távoktatást megelőző időszakban történt. (A vizsgált fogalmazások a 1. számú mellékletben bemutatott szövegalkotás-feladat

megoldásai, melynek témája a nagyszülőkkel való együttélés kérdése.) A kiválasztott 30 dolgozat ugyanazon budapesti általános iskola diákjainak munkája. A válogatás az alábbi szempontok szerint történt:

- Legyen összefüggő, értelmezhető a szöveg.
- A fogalmazás tartalma feleljen meg a feladatban megszabott témának.
- A dolgozat tartalmazzon legalább három érvet.¹

Fontos leszögezni, hogy a felvételi javítókulcs korábban említett terjedelmi kritériuma a válogatásban nem játszott szerepet, így terjedelem szempontjából nagyon eltérő szövegek is bekerültek a mintába. A kutatás anyagának teljes terjedelme megközelítőleg 3500 szövegszó. Az anyag előkészítése során az eredetileg kézzel írt dolgozatok gépelt változatát helyesírási szempontból javítottam a könnyebb feldolgozás érdekében, illetve az eredeti feladatlapon esetlegesen szereplő tartalmi vázlatokat, valamint a szövegekben megjelenő önjavításokat (áthúzott szavakat, mondatokat) az elemzett gépelt átiratokban nem tüntettem fel, csak a tanuló által véglegesnek tekintett szövegváltozatot vettem figyelembe. Más (például nyelvhelyességi) szempontból azonban nem történt változtatás

2. 2. Első fázis (saját korpusz)

A kutatás első fázisában az e-magyar elemzőlánccal készített annotáció alapján listáztam a fogalmazásokban található kötőszavakat. Az automatikus elemzést manuálisan ellenőriztem, a tévesen jelölt szóalakok címkézését javítottam. A további lépéseket erre a javított annotációra építettem. Az annotáció második lépésében kijelöltem az oksági viszonyt kifejezőket a kötőszavak közül. Ezt a lépést manuálisan végeztem, így minden esetben figyelembe vettem a szöveghelyeztetet a kategorizációkor. Rögzítettem az egyes kötőszókra vonatkozó gyakorisági adatokat a teljes anyagban (egyrészt az összes előfordulást, másrészt az oksági viszonyt kifejező előfordulásokat tekintve), emellett megjelöltem a bizonytalan eseteket.

2. 3. Második fázis (MNSZ)

A kutatás második fázisában arra a kérdésre kerestem a választ, hogy milyen hasonlóságok és eltérések figyelhetők meg a kötőszók eloszlásában a vizsgált diákesszékben, illetve a referencia korpuszban. Referenciaként a Magyar Nemzeti Szövegtár (MNSZ) beszélnyelvi, hivatalos, sajtó, személyes-fórum és személyes-közösségi alkorpuszait vizsgáltam együttesen (a továbbiakban erre az alkorpuszra MNSZ-H megnevezéssel hivatkozom), valamint a tudományos alkorpusz az előbbiektől elkülönítve (a továbbiakban MNSZ-T), tekintettel arra, hogy a szaknyelvi sajátosságok stiláris jellemzőket, így a kötőszóhasználatot is

¹ Az utolsó szempontot azt hivatott biztosítani, hogy a vizsgálatba bevont dolgozatokban ténylegesen szerepeljenek ok-okozati viszonyok, azonban ez a válogatási elv azt is eredményezte, hogy a mintába jellemzően a jobban teljesítő diákok munkái kerültek be. A későbbiekben a kutatás tanulságaira építve a vizsgálat kiterjeszthető nagyobb mintára, amely alapján általánosabb képet kaphatunk a korosztály nyelvhelyességéről.

befolyásolhatják. (A szépirodalmi alkorpuszt a hétköznapi nyelvhasználattól esetlegesen eltérő nyelvi jellemzők miatt nem vontam be a vizsgálatba.) Régió szerint mindkét esetben csak a magyarországi szövegekben kerestem, mivel fogalmazásokat budapesti diákok írták. A lekérdezéseket diákok dolgozataiból nyert kötszólista alapján végeztem.

Első lépésként a fogalmazásokban oksági viszonyt kifejező kötszók listájának elemeire kerestem a fent jellemzett két alkorpuszban, majd a fogalmazások annotálásakor kétes besorolásúnak bizonyult kötszók (*ha, hogy, hogyha*) esetében az MNSZ felületén generált véletlen mintán végeztem ellenőrzést. A minták elemszámát minden esetben 10%-os konfidenciaintervallumnak megfelelően szabtam meg.

A vizsgálat két fázisában nyert adatokat statisztikai módszerekkel (relatív gyakoriságok, eloszlások) elemeztem, vettem össze.

3. Eredmények

3. 1. Az annotáció során felmerült kérdések

Az e-magyar elemzőlánc annotációja alapján a 30 fogalmazás összesen 427 kötszót tartalmazott. Ezek közül az ellenőrzés során 7 bizonyult téves találatnak: Például a *vagy* egyes szám második személyű igealakot több esetben címkézte a vele azonos alakú kötszóként az elemző, illetve az *azért* szóalakot azokban az esetekben címkézte hibásan a program, amelyekben partikulaként szerepelt a mondatban, például:

(1) *Hiszen azért a szülők is szeretnek időnként kettesben lenni [...]*.

A szövegben egyértelműen kötszói szerepet betöltő *ezáltal* és *ezért* szóalakok összesen három esetben voltak hibásan, nem kötszóként címkézve. A fentiek alapján az annotáció pontossága 0,98, fedése 0,99, a hibák aránya tehát elenyésző.

Az annotáció második, manuálisan végrehajtott lépésében a szövegkörnyezetet figyelembe véve jelöltem meg az oksági viszonyt kifejező kötszavakat. A kategorizáció elsődleges szempontja a szemantikai, logikai tartalom milyensége volt, a tagmondatok közötti grammatikai kapcsolat típusát segédszempontként, egyes bizonytalan esetek osztályozásakor vontam be.

A 430 kötszó közül 123-at jelöltem meg oksági viszont kifejezőként, további 22 darabot pedig bizonytalan kategóriájúként, utóbbi típusba négy szóalak: a *hogy*, a *hogyha*, a *ha* és az *és* egyes előfordulásai kerültek. A *hogy*, a *hogyha* és a *ha* számos esetben vezetett be olyan következményes vagy feltételes mellékmondatot, amelyekben egyértelműen jelen van az ok-okozati viszony, bizonyos mondatok esetében azonban kérdéses a besorolásuk. Két tipikusan problémás eset egyike a beszélt nyelvben vélhetően megszokott, de pontatlan kötszóhasználat. A következő mondat ugyan nagyon hasonlít a feltételes mellékmondatot tartalmazókra, azonban az ezekre jellemző *ha ... akkor* kapcsolás

helyett használt *ha ... addig* arra is utalhat, hogy a *ha* itt időhatározói alárendelésben, az *amíg* helyett, tévesen áll az első tagmondat elején:

(2) [...] ha a szüleink elmennek valahova [...], *addig el tudunk hozzájuk utazni* [...].

Szintén a pontatlan vagy hiányos kötőszóhasználat okoz bizonytalanságot a következő mondat és kötőszavának besorolását illetően:

(3) *Ők az én családom, és [ezért] mindenről joguk van tudni.*

A másik kérdéses típust azok a mondatok alkotják, amelyeknek az első tagmondata az érvelő szövegben gyakran megjelenő minősítő vagy álláspontjelölő funkciót tölt be. Ezek a tagmondatok utalnak a feladat szövegében felvetett álláspontok, alátámasztandó tételek egyikére (Jobb a nagyszülőkkel együtt lakni. / Jobb a nagyszülőktől külön, csak a szülőkkel, testvérekkel élni.), azonban explicit módon nem szerepel bennük az éppen alátámasztani vagy cáfolni kívánt állítás. A feladat és a tanuló által alkotott szöveg kontextusában ezeknek a mondatoknak van oksági jelenéstartalma (hiszen érveket vagy ellenérveket fogalmaznak meg), azonban a szövegből kiragadott mondatok tagmondatai között az ok-okozati összefüggés az okozat hiánya miatt elhalványul, bár az első tagmondatban szereplő *előny, hátrány, érv, ellenérv, jó, rossz* stb. kifejezések továbbra is felidéznek az érvelő szövegekre jellemző sémákat. Erre a típusra példák az alábbiak:

(4) *A másik érvem az lenne, hogy szinte mindig lesz valaki a házban, így soha nem leszel magányos.*

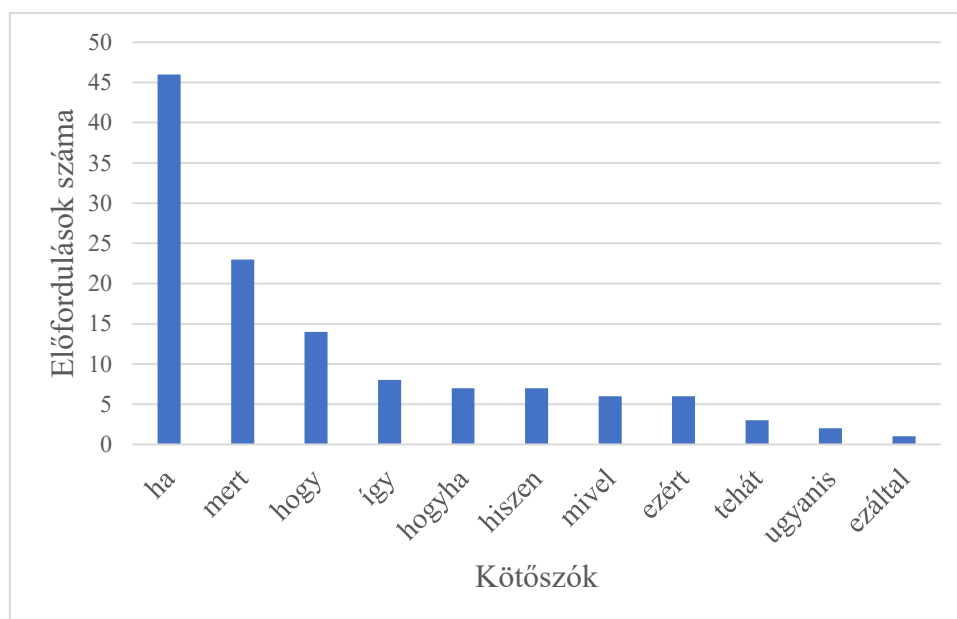
(5) *Persze jól jöhet, ha például a nagyszüleinket megkérdezzük, milyen volt a háborúban harcolni.*

(6) *Ami szerintem hátrányos az az, hogy így a házimunka nehezebb lesz [...]*

Az említett 22 bizonytalan esetet jelen dolgozat keretében nem vizsgálom, ezek esetében a kategóriák kritériumainak pontosítására, további grammatikai és szövegtani alapozásra, illetve lehetőség szerint társzkodolók bevonására lesz szükség.

3. 2. Az ok-okozati kapcsolatot jelölő kötőszók eloszlása a fogalmazásokban

A 123 oksági viszonyt jelölő kötőszó-előfordulás összesen 11 típus között oszlik meg, azonban gyakoriságuk nagyon eltérő. Az oksági viszonyt jelölőként kategorizált kötőszó eloszlását a 2. ábra szemlélteti.

2. ábra: Az oksági viszonyt jelölő kötőszók összesített eloszlása a fogalmazásokban

A leggyakoribb a feltételes mellékmondatot bevezető *ha*, amely összesen 46-szor, azaz az esetek több mint harmadában fordult elő. Ezt követi gyakoriságban a 23-szor előforduló *mert*, majd a 14-szer előforduló, jellemzően következményes mellékmondatot bevezető *hogy*. A további kilenc kötőszó mindegyike 10-nél kevesebbszer, az esetek kevesebb mint 6 százalékában jelent meg a szövegek összességében.

A fogalmazásokban oksági viszonyt kifejező kötőszókat az e-magyar (három korábban említett előfordulás esetében korrigált) annotációjában megjelölt összes kötőszóval összevetve azt látjuk, hogy a fenti diagramon szereplő 11 típus közül 8 minden előfordulásakor oksági viszonyt jelöl, mindössze három típus: a *ha*, a *hogy* és a *hogyha* tér el ettől. Az oksági jelentéstartalmú kapcsolatot kifejező és az összes előfordulás aránya a *ha* esetében 73%, a *hogy* esetében 23%, a *hogyha* esetében pedig 70%. (Fontos azonban megemlíteni, hogy éppen ez a három kötőszó jelenti a 22 bizonytalan eset zömét, tehát ezek az arányok a kategorizáció pontosítása után valamivel nagyobbak is lehetnek.) Eszerint azt feltételezhetjük, hogy a *ha* és a *hogyha* gyakran vesz részt oksági viszony kifejezésében, a *hogy* azonban előfordulásainak számához viszonyítva ritkábban, más jellegű kötőszói funkciói gyakoribbak lehetnek.

A bevezetés 1.4. fejezetében az 1. számú táblázat foglalja össze a Magyar grammatika vonatkozó fejezetei alapján az ok-okozati viszony egyes altípusait kifejező kötőszókat. A dolgozatokban előforduló kötőszók ezen altípusok szerint összesített gyakoriságát a 2. táblázat mutatja be.

3. táblázat: Az oksági kötőszók altípus szerinti eloszlása a vizsgált érvelő fogalmazásokban

Funkció szerinti típus	Kötőszó	Előfordulások száma (kötőszó)	Előfordulások száma (funkció)
Ok-okozat	<i>így</i>	8	16
	<i>ezért</i>	4	
	<i>tehát</i>	3	
	<i>ezáltal</i>	1	
Okozat-ok	<i>mert</i>	23	38
	<i>hiszen</i>	7	
	<i>mivel</i>	6	
	<i>ugyanis</i>	2	
Feltételes mellékmondat	<i>ha</i>	46	53
	<i>hogyha</i>	7	
Következtető mellékmondat	<i>hogy</i>	14	14

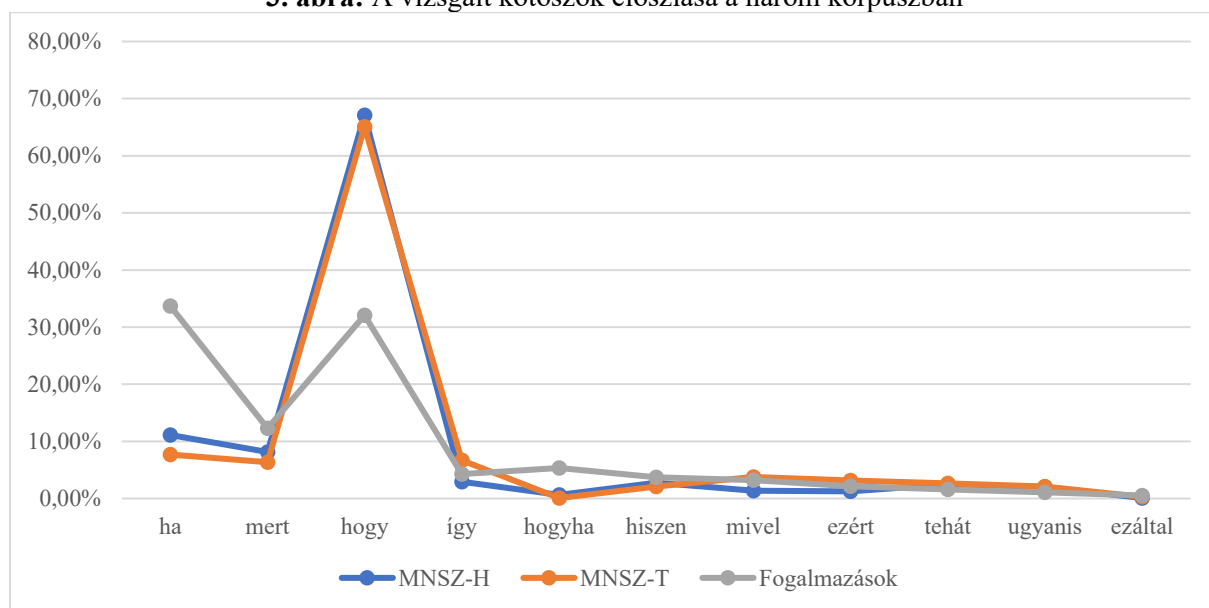
Az összesítésből kitűnik, hogy a feltételes jelentéstartalmú tagmondatok majdnem az esetek felét teszik ki, ezt követi az okozat-ok sorrendű alá- vagy mellérendelő viszonyban lévő tagmondatokat összekapcsoló típus, míg a deduktív gondolatmenetre utaló ok-okozat sorrendű összekapcsolást, illetve a következtető mellékmondatot bevezető kötőszók csak 30 esetben szerepeltek a dolgozatokban.

4. A vizsgált kötőszók gyakorisága a reprezentatív korpuszban

A vizsgálat második fázisában arra a kérdésre kerestem a választ, hogy milyen hasonlóságok és eltérések figyelhetők meg a vizsgált kötőszók eloszlásában a kutatás első fázisának anyagát alkotó diákesszékből, illetve az MNSZ referenciakorpuszként kijelölt alkorpuszaiban.

A 3. ábra a diákesszék elemzésekor kijelölt 11 kötőszó összes előfordulásának eloszlását mutatja a három vizsgált korpuszban.

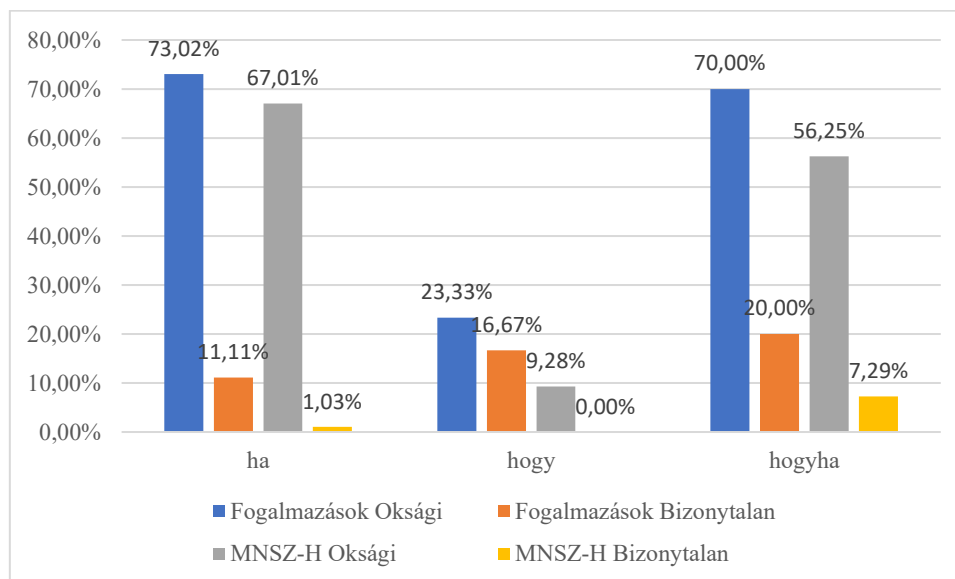
3. ábra: A vizsgált kötőszók eloszlása a három korpuszban



A legnagyobb eltérés a *ha* és a *hogy* kötőszók esetében látható: a *ha* a fogalmazásokban a másik két korpuszhoz képest több mint háromszor gyakoribb, míg a *hogy* kevesebb mint feleakkora arányban fordul elő a dolgozatokban. Bár az MNSZ két alkorpuszának adatai között kisebbek a különbségek, a khi-négyzet próba alapján a két minta nem tekinthető azonos eloszlásúnak ($p < 10^{-7}$), azonban a fogalmazások adataival összevetve így is jóval kisebb az eltérés az MNSZ két alkorpusza között, ezért a továbbiakban az MNSZ-H korpusz és a fogalmazáskorpusz elemzését vetem össze.

Azon három kötőszó előfordulásainak funkcióit, amely a fogalmazások elemzése alapján nem minden esetben oksági viszonyt jelöl, az MNSZ-H korpuszból generált véletlen mintán ellenőriztem. A minta elemszámának meghatározásakor 10%-os konfidenciaintervallumot választottam, így a *ha* és a *hogy* esetében 97, a *hogyha* esetében pedig 96 előfordulást vizsgáltam. Az ellenőrzés eredményei (az oksági viszonyt kifejező, illetve a bizonytalan kategorizációjú elemek és az adott mintában szereplő teljes elemszám aránya) a fogalmazások annotálásának tapasztalataival összevetve az 4. ábrán láthatók.

4. ábra: A *ha*, *hogy* és *hogyha* kötőszókra vonatkozó véletlen minta kézi annotálásának eredményei



Az MNSZ-ből nyert véletlen minta esetében kisebbnek tűnik a bizonytalan esetek aránya, mint a fogalmazásokat tartalmazó korpuszban, ugyanakkor ez nem feltétlenül a tanulók nyelvhasználatát, kötőszóhasználatának pontosságát minősíti, hanem ezt az is okozhatja, hogy az MNSZ-ből generált véletlen mintánál kisebb abszolút gyakorisággal jelentek meg az egyes kötőszók a korpuszban (a *hogyha* mindössze tízszer). Az MNSZ adatait tekintve a *hogyha* előfordulásainak címkézése eredményezte a legtöbb bizonytalan besorolást: ezek aránya 7,23% szemben a másik két kötőszó 1,03%-os, illetve 0%-os mutatójával. Ezeknek a bizonytalan esetek többsége beszélt nyelvi szövegek átiratából származott. Ilyen esetre példa az alábbi két (7. és 8. számú) mondat:

(7) *Az úriembert, akit leültettek, természetesen semmi nincsen a nevén, tehát, hogyha én, amit nagyon nem szeretnék, de hogyha polgári perre kerül a sor, tehát, hogyha elveszik tőlünk a kocsit, mert jelen pillanatban perelni sem tudunk, mert igazából jogilag mienk az autó, tehát nincs mire perelni, de hogyha polgári perre kerül a sor [...]*

(8) *Dr. Kelényi Lóránt agykutató professzor: - Kaptak egy-egy fejhallgatót, lefekszik, becsukja a szemét, kezében van egy nyomókönyv, különböző trükkös összeállítású hangokat kap kis füttyöket, amely füttyöket ő neki, ami hogyha megismerte, meg kell nyomni a gombot, és hogyha ezt sorozatban elrejtve valami van, fel kell ismernie, amit mondtam, hogy most azt ismerje fel, és közben fejszámolást végez, hogy ez hanyadik volt és így tovább [...]*

Szintén a korpusz szövegeinek sajátossága, hogy nem minden szöveg hibátlan helyesírási szempontból, fórumhozzászólásokban éppen a kötőszavak különvagy egybeírása gyakori probléma:

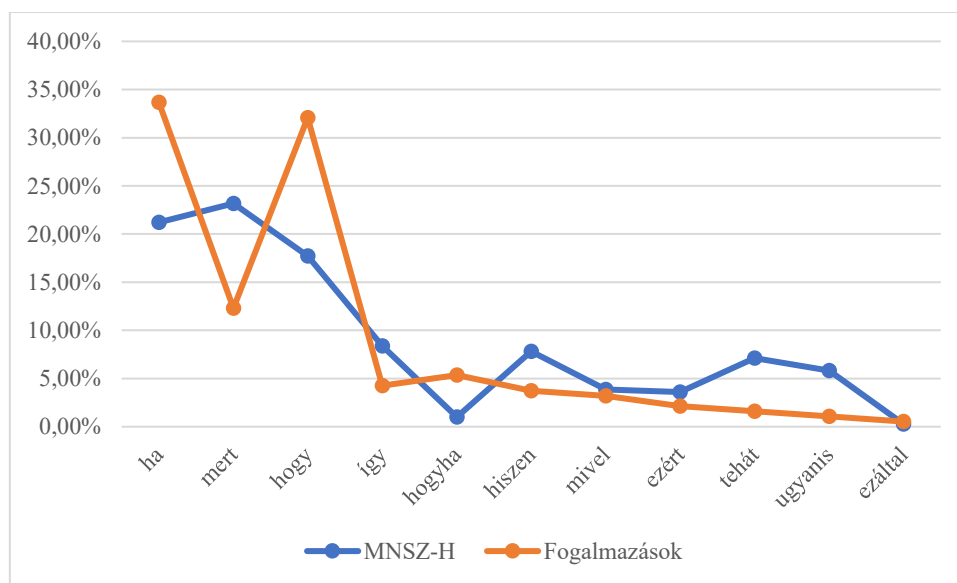
(9) Szóba került az is, hogyha az idős ember nincs otthon, akkor bezárja a lakás külső, udvari ajtaját.

Ezek ugyan egyértelműen kategorizálható esetek, de egy későbbi, kiterjedtebb elemzés esetén fontos figyelmeztetés lehet az annotátorok számára.

A véletlen minta manuális szűrésének eredményeit a három vizsgált kötőszó találati számira vetítve a *ha* relatív gyakorisága 11,14%-ról 21,22%-ra; a *hogy*-é 67,12%-ról 17,71%-ra; a *hogyha* relatív gyakorisága pedig 0,63%-ról 1,01%-ra módosul.

Az így korrigált arányokat a fogalmazásokból nyert eloszlással összevetve (lásd 5. ábra) továbbra is érvényes, hogy a *ha* kötőszó oksági viszonyt kifejező szerepben a dolgozatokban sokkal gyakoribb, mint az a reprezentatív korpusz alapján várható lenne (ez lehet anyanyelv-elsajátítás folyamatával összefüggő sajátosság, de az érvelő szövegtípusból is adódhat), hasonló különbség látható a *hogy*, és kisebb mértékben, de szintén a fogalmazások javára a *hogyha* kötőszó gyakoriságát tekintve is.

5. ábra: Az MNSZ-H korrigált eloszlása és a fogalmazásokból nyert eloszlás



A többi kötőszó esetében azonban kivétel nélkül az MNSZ alkorpuszában nagyobb a relatív gyakoriság: az eredmények azt mutatják, hogy a vizsgált fogalmazások az MNSZ alkorpuszában szövegeihez képest kevésbé változatosak a kötőszóhasználat szempontjából.

5. Összegzés, következtetések

A 30 fogalmazás elemzése rámutatott arra, hogy a diákok fogalmazásaiban az oksági viszonyt kifejező kötőszók között kiugróan nagy arányban fordulnak elő a funkciójukat tekintve feltételes mellékmondatot bevezető kötőszók (különösen a *ha*), míg a deduktív gondolatmenetre utaló ok-okozat sorrendű tagmondatokat

összekapcsoló kötőszók aránya igen kicsi, ám ilyen funkcióban több különböző kötőszó fordult elő a szövegekben. Felvetődik a kérdés, hogy ennek az eltérésnek a magyarázata a tanulók érvelési és szövegalkotási képességének fejlettségében, vagy az őket tanulmányaik során körülvevő szövegkorpusz mint nyelvi minta sajátosságaiban keresendő.

Az elemzés egyik fontos tapasztalata, hogy az beszélt nyelvi alkorpuszt érdemes külön kezelni, vagy kizárni az összetett mondatok szerkezetét érintő vizsgálatok esetén, mert a spontán beszéd alacsonyabb szerkesztettsége, az utólagos önellenőrzés és -korrekció lehetőségének hiánya gyakran eredményez befejezetlen, vagy kevésbé összefüggő mondatokat, amelyekben a grammatikai szerkezetből adódóan fennáll a lehetősége az oksági kapcsolatnak, azonban a következmény hiányzik, „elvész” a gondolatfolyam további láncszemei között.

A kutatás anyagát alkotó érvelő fogalmazásokban megfigyelt kötőszóhasználat lényegesen eltér az MNSZ választott alkorpuszaiból nyert eloszlásoktól, ennek a megfigyelésnek az alaposabb alátámasztásához, okainak feltárásához azonban szükséges lenne a vizsgálatot rétegzett mintavétellel, több résztvevővel (illetve lehetőleg az egyes diákok több, esetleg különböző műfajú fogalmazásával) és több annotátorral megismételni. Jelen dolgozatban a tudományos alkorpusz adatainak részletesebb elemzésére nem került sor, azonban a vizsgálat folytatásaként célszerű lenne ezt is elvégezni, valamint bővíteni a referenciakorpuszok körét. Bár a korosztály nyelvhasználatára nézve hasznos adalékként szolgálhatna például a Gyermeknyelvi Beszédatadtbázis (GABI), a beszélnyelvi átiratokkal kapcsolatban fent leírt tapasztalatok alapján ezt célszerű külön kezelni, esetleg a korosztály beszélt és írott nyelvi jellemzőinek összevetése esetén lehetne alkalmazni. Az érvelő szövegek sajátosságainak vizsgálata szempontjából pedig érdemes lehet ilyen típusú szövegeket nagyobb arányban tartalmazó korpuszt bevonni, erre alkalmas lehet például egy hétköznapi nyelvhasználathoz közelebb álló argumentatív szövegekből, például termékértékelésekből összeállított korpusz.

Az írott szövegekben megjelenő oksági viszonyt kifejező kötőszók, illetve más hasonló funkciójú nyelvi elemek vizsgálatának eredményei, kategóriarendszerének kidolgozása alkalmazásának lehetséges területei többek között a szakdidaktika és a pedagógiai értékelés, az alkalmazott retorika és a nyelvtechnológiai fejlesztések, illetve az ilyen irányú kutatások alapozó lépései lehetnek egy argumentatív szövegekből épített korpusz annotációjának.

Irodalom

- Adamik Tamás, A. Jászó Anna & Aczél Petra (2005). *Retorika*. Budapest: Osiris Kiadó.
- Arisztotelész (1982). *Rétorika*. Adamik Tamás ford. Budapest: Gondolat.
- Bánki Dezső 2004. Az érvelés mestersége. *Budapesti Könyvszemle*, 3, 272–276.
- Becker-Mrotzek, M. & Böttcher, I. (2014). *Schreibkompetenz entwickeln und beurteilen*. Cornelsen Berlin: Schulverlage GmbH.
- Bereiter, C. & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale: L. Erlbaum Associates.

- Csákberényiné Tóth Kinga** (2011). Fantáziafogalmazások jellemzői 4. osztályban. *Anyanyelv-pedagógia 1*. Letöltés: <https://www.anyanyelv-pedagogia.hu/cikkek.php?id=301>
- Deme László** (1965). A mondatok egymáshoz kapcsolódása a beszédben. *Magyar Nyelvőr*, 89, 292–302.
- e-magyar** = Váradi Tamás, Simon Eszter, Sass Bálint, Gerócs Mátyás, Mittelholcz István, Novák Attila, Indig Balázs, Prószéky Gábor & Vincze Veronika (2017). Az e-magyar digitális nyelvfeldolgozó rendszer. In Vincze Veronika (szerk.), *XIII. Magyar Számítógépes Nyelvészeti Konferencia. MSZNY 2017* (49–60). Szeged: Szegedi Tudományegyetem, Informatikai Intézet.
- GABI** = Bóna Judit (2017). GABI – Gyermeknyelvi beszédatadtbázis a kutatásban In Bóna Judit (szerk.), *Új utak a gyermeknyelvi kutatásokban* (35–50). Budapest: ELTE Eötvös Kiadó.
- Horváth Zsuzsanna** (1998). *Anyanyelvi tudástérkép. Középiskolai tantárgyi feladatbankok III.* Budapest: Országos Köznevelési Intézet.
- Horváth Zsuzsanna** (2008). Írni jó. A szövegalkotás fejlesztése. In Bánkuti Zsuzsa & Lukács Judit (szerk.), *Tanulmányok az érettségiről: hatásvizsgálat, tantárgyi vizsgák értékelése, feladatfejlesztés* (323–351). Budapest: Oktatókutatási és Fejlesztési Intézet.
- Javítókulcs 2024** = *Magyar nyelv a 8. évfolyamosok számára. Javítási-értékelési útmutató 2024. január 20.* Letöltés: https://www.oktatas.hu/pub_bin/dload/kozoktat/keiskolazas/feladatsorok/2024/A8_2024_1_ut.pdf
- Juhász Milán** (2021). Az érvelő esszé értékelési szempontjainak megítélése pályakezdő és gyakorlott pedagógusok által. *Anyanyelv-pedagógia 1*, 20–33. doi: 10.21030/anyp.2023.2.5
- Kádárné Fülöp Judit** (1990). *Hogyan írnak a tizenévesek: Az IEA fogalmazásvizsgálat Magyarországon.* Budapest: Akadémiai Kiadó.
- Major Hajnalka** (2015). *Retorika és szövegalkotás.* Budapest: Eötvös Kiadó.
- Margitay Tihamér** (2007). *Az érvelés mestersége.* Typotex.
- MG** = Keszler Borbála szerk. (2000) *Magyar grammatika* Budapest: Nemzeti Tankönyvkiadó.
- MNSZ** = Oravecz Csaba, Váradi Tamás & Sass Bálint (2014). The Hungarian Gigaword Corpus. In: *Proceedings of LREC 2014.*
- Molnár Edit Katalin** (2000). A fogalmazási képesség fejlődésének mérése. *Iskolakultúra*, 8, 49–59. Letöltés: <https://www.iskolakultura.hu/index.php/iskolakultura/article/view/19315>
- Nagy Zsuzsanna** (2013). A fogalmazásértékelés megbízhatósága két független bíráló értékítéleteinek elemzése alapján. *Magyar Pedagógia*, 3, 153–179. Letöltés: http://www.magyarpedagogia.hu/document/2_Nagy_MP1133.pdf
- NAT 2012** = Nemzeti Alaptanterv (2012) *Magyar Közlöny* 66, 10639–10847.
- NAT 2020** = Nemzeti Alaptanterv (2020) *Magyar Közlöny* 17, 290–446.
- OH 2022** = *Magyar nyelvi feladatlap a 8. évfolyamosok számára: 2022. január 22. 10:00 óra* Letöltés: https://www.oktatas.hu/pub_bin/dload/kozoktat/keiskolazas/feladatsorok/2022/A8_2022_1_fl.pdf
- Orosz Sándor** (1972). *A fogalmazástechnika mérés-metodikai problémái és országos színvonala.* Budapest: Tankönyvkiadó.
- P. Tóth Teodóra** (megj. a.). *Az érvelő esszé a középiskolai felvételin.* In Bajzát Tímea Borbála, Hauber Kitti, Molnár Emese, Kopcsák Róbert, Mayer Péter & Turcsik Biborka (szerk.), *Nyelvek és nézőpontok. Tanulmányok a 17. és 18. Félúton konferenciáról* (76–89.). Budapest: Eötvös Loránd Tudományegyetem Bölcsészettudományi Kar, Nyelvtudományi Doktori Iskola.
- SRDP 2018** = *Az SRDP oktatási nyelvi szövegtípusainak katalógusa.* (2018). Bundesministerium Bildung, Wissenschaft und Forschung.
- Szikszainé Nagy Irma** (2006). *Leíró magyar szövegtan.* Budapest: Osiris Kiadó.
- Tájékoztató 2023** = *Tájékoztató a középfokú beiskolázás egységes írásbeli felvételi vizsgáinak - magyar nyelvi és matematika - feladatlapjairól (2023),* Oktatási Hivatal. Letöltés: https://www.oktatas.hu/koznevel/kozepfoku_felveteli_eljaras/tajekoztato_matek_magyar
- Tolnai Varga Piroska** (2018). Diskurzusszervező elemek a diákesszéekben. *Anyanyelv-pedagógia 1*, 28–39. doi: 10.21030/anyp.2018.1.3
- Tóth László** (2006). *Az írásbeli szövegalkotás pszichológiai alapjai.* Debrecen: Pedellus Tankönyvkiadó.

Melléklet

1. számú melléklet: Az adatgyűjtéskor használt feladatsor szövegalkotás-feladata

Egyes családokban a gyerekek nem csak szüleikkel és testvéreikkel, hanem nagyszüleikkel is egy háztartásban élnek. Szerinted az a jobb, ha több generáció lakik együtt, vagy ha a gyerekek csak a testvéreikkel és a szüleikkel élnek egy fedél alatt?

10–12 mondatos fogalmazásodban írd legalább 2-2 érvet mindkét lehetőség mellett, és fejtsd ki, hogy számodra melyik álláspont az elfogadhatóbb!

Figyelem! A 4 mondatnál rövidebb fogalmazás nem értékelhető! Ha a fogalmazás nem éri el a 10 mondatot, az jelentős pontlevonással jár.

.....

.....

.....

.....