

Kutatási terv

Gazdik Anna

2011. június 22.

1. A kutatás célja

A kutatásom célja a magyar mondat szerkezetének vizsgálata és modelljének kidolgozása a Lexikai-Funkcionális Grammatika (LFG) keretében. Az elemzés a szintaxison túlmenően kiterjed az információs szerkezetre, a diskurzus szintjére, és ezek összefüggéseire a szintaktikai szerkezettel. Az LFG, mint nem transzformációs, moduláris elmélet lehetővé teszi az eltérő nyelvi információk disszociációját és különböző nyelvi szinteken történő ábrázolását. A kutatás hozzájárulna mind a magyar mondat szerkezet kutatásához, mind pedig az LFG, mint elméleti keret fejlődéséhez. Az előbbi szempontjából érdemes kiemelni, hogy a magyar mondat elemzésére eddig főként a generatív transzformációs modellben (lásd például É. Kiss (2002)) és HPSG-ben (Szécsényi, 2011) születtek javaslatok, így mindenképpen érdemes megvizsgálni, mit tud hozzátenni ezekhez az LFG-s megközelítés. Másfelől, az LFG modelljében mind az információs, mind pedig a diskurzus szerkezet aktuális kutatási területnek számít, amelyhez nagyban hozzájárulnának a magyar adatokon végzett vizsgálatok. A továbbiakban részletesebben ismertetem a kutatási területet, az előzményeket és a főbb megoldandó problémákat.

2. A kutatás előzményei

Doktori disszertációmiban a magyar és francia többszörös kérdéseket vizsgáltam és elemeztem a Lexikai-Funkcionális Grammatika keretében. Azonban mind a többszörös kérdések, mind pedig a hozzájuk kapcsolódó magyar mondattani jelenségek vizsgálata túlmutat a doktori disszertáció keretein, és további kutatásokat igényel.

A többszörös kérdések elemzése során hamar világossá vált, hogy pusztán a szintaxis figyelembe vétele nem teszi lehetővé, hogy számot adjunk ezen szerkezetek minden tulajdonságáról, főként az értelmezési lehetőségeikről. Mind a franciában, mind a magyarban találunk olyan szerkezeteket, amelyeket párlistával és specifikus párral is meg lehet válaszolni:

- (1) **Qu'est-ce que** tu as donné **à qui** ?
mit te AUX ad.PART kinek
Mit adtál kinek?
- (2) **Ki** vesz el **kit**?
- (3) **Hol** és **hogyan** töltötte a nyarakat gyerekkorában?

Az (3)-as példában ráadásul a párlista olvasat nem is magukból a kérdőszókból, vagy azok szintaktikai pozíciójából következik, hanem a többes számú tárgyból: a kérdés minden gyermekkori nyárral kapcsolatban választ vár a *hol* és *hogyan* kérdőszókra.

Ezek a megfigyelések arra engedtek következtetni, hogy a szintaxison kívül a mondat információs szerkezetét, továbbá a mondat szintjén túl az adott diskurzust is figyelembe kell venni az elemzésben.

A fősodratú LFG-ben általában használt információs szerkezet a következőképpen épül fel: négy halmazból áll (topik, fókusz, háttér és kiegészítő információ), amelyeket két jegy kombinációival definiálnak (új és prominens) a következőképpen (Butt and King, 1996):

	Topic	Focus	Background Inf.	Completive Inf.
New	–	+	–	+
Prominent	+	+	–	–

Amennyiben ezt a szerkezetet vesszük alapul, ebben kell elhelyeznünk a többszörös kérdés kérdőszavait, és ebből levezetni az értelmezésbeli különbségeket is (azaz, hogy a kérdésre párlistával vagy specifikus párral válaszolunk). Szintaktikai, szemantikai, prozódiai és morfológiai érvek egyaránt amellett szólnak, hogy a párlista választ váró kérdésekben legalább az egyik kérdőszó a *TOPIK* halmazban helyezkedik el. Egy lehetséges megközelítés szerint ebből adódik maga a párlista olvasat: a topik halmazban megjelenő kérdőszó által denotált (kontextuálisan meghatározott, vagy ismert) halmaz leképeződik a másik kérdőszó által denotált halmazra. A specifikus pár választ váró kérdések mindkét kérdőszava a *FÓKUSZ* halmaz része, nem jön létre a leképezés a kettő között.

Azonban nem minden magyar mondattani jelenség ábrázolható ezzel az információs szerkezettel. Tekintsük a következő példákat:

(4) Ma a **feleségem** vitte az óvodába a gyerekeket.

A (4) példában a *feleségem* összetevő, bár a preverbális pozícióban van, sem prozódiailag, sem szemantikailag nem tekinthető fókuszsnak. Az ilyen összetevőket *hókuszsnak* nevezi a szakirodalom és úgy határozza meg őket, mint olyan azonosító funkciójú elemek, amelyek általában egy rendszeres esemény egy nem szokványos, valamiért meglepő szereplőjét azonosítják (Kálmán, 2001). Kérdés, hogy a fentebb bemutatott struktúrában hol lenne a hókusz helye. Hasonlítsuk most össze a két következő példát:

(5) [_{Top}János] [_Qmindig] [_{Foc}MARIVAL] találkozik a városban.

(6) [_{CTop}Ritkán] [_{Foc}JÁNOS] jár moziba.

A fenti példákban a mondatkezdő összetevők mind formailag (prozódia), mind szemantikájukban eltérnek egymástól. -ben a mondatkezdő összetevő (a magyar nyelvű szakirodalom terminusával) *topik*, míg a *d*-ben *kontrasztív topik*. Anélkül, hogy belemennénk a részletekbe, látható, hogy a két összetevő eltér (többféle elem jelenhet meg kontrasztív topikként, mint hagyományos topikként, illetve használatuk más diskurzusmintákhoz köthető), így ezeknél a példáknál sem világos, hogyan lehetne mindkétfajta topikot (a hagyományos topikokat célszerűbb *thematic shifternek* nevezni) az információs szerkezet *TOPIK* halmazába utalni, hiszen ez egybemosná a köztük fennálló különbségeket, és nagyban megnehezítené a *TOPIK* információszerkezeti halmaz definiálását.

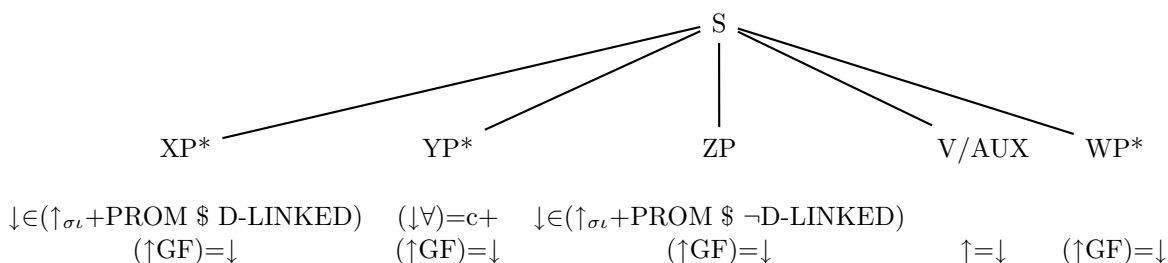
A magyar mondat szerkezet LFG-ben történő ábrázolása tehát egy olyan felépítésű információs szerkezetet tesz szükségessé, amely nem tartalmazza a topikot és a fókuszt alaphalmazokként, hanem elég általános ahhoz, hogy a különböző topikfajták, a fókusz, a hókusz, és a különböző

kérdőszók egyaránt helyet kapjanak benne. Egy ilyen i-struktúra épülhet ezen elemek közös szemantikai tulajdonságaira, például arra, hogy egyes elemek szemantikai és formai szempontból prominensek, míg mások nem, illetve egyes elemek a mondatot az adott diskurzushoz kapcsolják, míg mások nem. A két alaptulajdonság így a *prominencia* és a *D-linkedness*. Az információs szerkezet pedig így épülne fel:

$$\left[\begin{array}{l} +\text{PROM} \left[\begin{array}{l} \neg \text{D-LINKED} \{ \text{FOCUS, QW, HOCUS} \} \\ \text{D-LINKED} \{ \text{TH SHIFTER, CONTRASTIVE TOPIC, QW} \} \end{array} \right] \\ -\text{PROM} \left[\begin{array}{l} \neg \text{D-LINKED} \{ \text{COMPLETIVE INFORMATION} \} \\ \text{D-LINKED} \{ \text{BACKGROUND INFORMATION} \} \end{array} \right] \end{array} \right]$$

Az információs szerkezet alkategóriái ebben a felépítésben tágabb szemantikai értelmezéssel bírnak, de nem is céljuk, hogy pontosan meghatározzák az adott halmazban megjelenő elemek szemantikáját. Az adott elemek szemantikája az információs szerkezetben elfoglalt hely és az adott diskurzus függvénye (kérdés-válasz pár, korrekció, kontraszt, narratíva, stb.). A javasolt információs szerkezetnek egy olyan felépítése is elképzelhető, amelyben az említett tulajdonságok önálló halmazokat alkotnak, azaz nem épülnek egymásba. A szerkezet felépítésének pontos meghatározása azonban további kutatásokat igényel.

Az LFG-ben ennek megfelelően leképezés létesíthető a szintaktikai szerkezet és az információs szerkezet között a magyarban:



Az ábra azt illusztrálja, hogy a preverbális pozícióba +PROM és -D-LINKED elemek kerülnek, míg a mondatkezdő elemek szintén +PROM de D-LINKED. Az LFG-ben feltételezett c-struktúra a magyarban lapos szerkezetű, hiszen a lineáris precedencia kifejezi a hatóköri viszonyokat, míg az X-vonás elméletből ismert maximum kétfelé ágazás az LFG-ben nem kötelező.

A fentiekben a magyar mondat LFG-s reprezentációjának első lépéseit vázoltam fel az LFG keretében. A következőkben rátérek azokra a problémákra, amelyek feltétlenül további kutatásokat igényelnek.

3. További megoldandó problémák

3.1. A preverbális pozíció kérdései

A preverbális pozíció problematikája régóta kutatott téma a magyar nyelvészetben. A szintaktikai szerkezetben gyakran az egyik ott előforduló elemről *fókusz pozíciónak* nevezték. A kutatásom célja egy olyan szintaktikai szerkezet felvázolása, amely nemcsak a fókuszos, hanem

a hókuszos és egyéb semleges mondatokról is számot tud adni. Ez elsősorban azt jelenti, hogy a mondat szerkezet a diskurzus szempontjából semleges, azaz nem tartalmaz diskurzusfunkciókra utaló funkcionális projekciókat. Ez azonban nem jelenti azt, hogy a szintaktikai szerkezet független az információs szerkezettől és a diskurzustól a magyarban. Az LFG moduláris rendszerében leképezés létesíthető a szintaktikai pozíciók és az információs szerkezet egyes részegységei között. A pontos leképezés feltérképezése és formalizálása azonban további kutatásokat igényel. Kérdés például a preverbális mezőben megjelenő kvantorok információszerkezeti státusza, különösen, ha az adott kvantor szemantikailag fókuszra tekinthető:

- (7) a. Q: Hányan nézték meg a filmeket?
b. A: A Csillagok háborúját **mindenki** megnézte.

További probléma a mondatvégi pozíció információszerkezeti státuszának meghatározása, ahol a fókuszra, mint prominens elemén kívül kontrasztív topikok is előfordulhatnak:

- (8) Fókusz: A **lányok** nyerték meg tegnap a **kajakversenyt**, a **fiúk** pedig a **kenuver-senyt**.
(9) Kontrasztív topik:
a. Q: **Mit** tanultak idén a diákok?
b. A: -Az elsőévesek szintaxist, a másodévesek pedig szemantikát tanultak.
c. B: -Szintaxist tanultak az elsőévesek, és szemantikát a másodévesek.

Amennyiben az A válaszban az *elsőévesek* és *másodévesek* kontrasztív topiknak tekintendők, kontrasztív topikok lesznek a B válaszban is, holott a mondat végén helyezkednek el.

3.2. Fókusz és kérdőszók második előfordulása

A preverbális pozíció kérdésköréhez tartozik, ám túl is mutat azon az olyan mondatok szerkezetének meghatározására, amelyekben két prominens preverbális elemet találunk:

- (10) Q: -Jó, de JÁNOS mit evett? (Kálmán, 2001)
(11) Q: -**Ki** NEM A HAMLETET olvasta? (Kenesei, 2009)

A példák közös tulajdonsága, hogy nem lehetnek diskurzuskezdők, azaz vagy a kérdőszók, vagy pedig a fókusz már másodsorra fordul elő bennük. Érdekes, hogy míg két preverbális fókusz nem grammatikus a magyarban, egy kérdőszó és egy fókusz előfordulhat, mégpedig meghatározott sorrendben. Ezek az adatok felvetik egy opcionális második prominens preverbális pozíció létét a magyarban, amely nincs mindig jelen a struktúrában, csak akkor, ha egy fókusz vagy kérdőszó másodsorra fordul elő benne. A jelenség pontos leírásához szükség van a másodsorra megjelenő fókusz és kérdőszók pontos szemantikai leírására, a diskurzusban játszott szerepük vizsgálatára, és a szintaktikai szerkezetben elfoglalt pozíciójuk meghatározására.

3.3. A diskurzus szerkezet formalizálása

Az eddig bemutatott problémákból és adatokból is kitűnik, hogy gyakran hivatkozunk az egyes mondatnál nagyobb nyelvi egységre: a diskurzusra. A diskurzusra, mint nyelvi elemzési szintre szükség van a diskurzusfunkciók meghatározásánál, de a fókusz és a kérdőszók második megjelenését is csak egy adott diskurzuson belül van értelme tárgyalni. A diskurzus, mint önálló nyelvi szint többször felmerült az LFG-ben, de még nem szerepel az alapvető reprezentációs szintek között. King and Zaenen (2004) javaslata alapján a diskurzust fagráffal érdemes reprezentálni, amely tükrözi a mondatok közt fennálló diskurzusrelációkat. A diskurzusrelációk ontológiájának meghatározásakor a kiindulópont Asher and Lascarides (2003) *Segmented Discourse Representation Theory* (SDRT), amelynek pontos integrációja az LFG elméleti keretébe fontos jövőbeli kutatások feladata. Meglátásom szerint a mondatokat az információs szerkezeteik reprezentálják a diskurzus szerkezetben, amelyek aztán a javasolt fagráf formájában jelenítik meg a különböző diskurzusrelációkat.

4. A kutatás jelen állása és aktualitása

Mint említettem, mind az információs, mind pedig a diskurzus szerkezet az aktuális és elkövetkező LFG-s kutatások fontos részét képezi. Ehhez nagymértékben hozzájárulhatnak a magyar adatokon végzett kutatások, amelyek például egy a topik-fókusz alapútól eltérő, annál általánosabb információs szerkezet bevezetését indokolhatják. A kutatás és az eddigi eredmények fontosságát az is mutatja, hogy egy kivonatát elfogadták az idei nemzetközi LFG konferenciára. A kutatás az LFG elméleti keretén túl is jelentőséggel bír: a magyar szintaxis és diskurzus összefüggéseiről szóló rész előadásra kerül az idei HPSG konferencia előtt megrendezett információszerkezeti workshopon és a szeptemberi *Constraints in Discourse* konferencián. Ugyanakkor magyar nyelvészeti kutatásokra is hatással lehet egy LFG alapú modell létrejötte, amely kiegészíti a generatív transzformációs modellekben és a HPSG-ben javasolt modelleket. Mindezekért is különösen fontosnak tartanám az MTA Nyelvtudományi Intézetében végzett kutatómunkát egy 3 éves tudományos kutatói állás keretében, ahol lehetőségem nyílna arra, hogy a magyar szintaktikai kutatások jelentős képviselőivel dolgozzam együtt, és hogy több modell megvitatásával és összehasonlításával egy teljesebb képet kapjunk a magyar mondat szerkezetéről.

Hivatkozások

- Asher, N. and Lascarides, A. (2003). *Logics of Conversation*. Studies in Natural Language Processing. Cambridge University Press, 2005 edition.
- Butt, M. and King, T. H. (1996). Structural topic and focus without movement. In Miriam, B. and King, T. H., editors, *Proceedings of the LFG96 Conference*, Rank Xerox, Grenoble. CSLI.
- É. Kiss, K. (2002). *The Syntax of Hungarian*. Cambridge Syntax Guides. Cambridge University Press, Cambridge.
- Kálmán, L., editor (2001). *Magyar leíró nyelvtan. Mondattan 1. [Descriptive Grammar of Hungarian. Syntax 1.]*. Tinta Könyvkiadó, Budapest.

- Kenesei, I. (2009). Quantifiers, negation, and focus on the left periphery in hungarian. *Lingua*, 119:564–591.
- King, T. H. and Zaenen, A. (2004). F-structures, information structure, and discourse structure. Proceedings of LFG04, extended abstract.
- Szécsényi, T. (2011). Magyar mondatszerkezeti jelenségek elemzése HPSG-ben. In Huba, B., editor, *Általános nyelvészeti tanulmányok*, volume 23., pages 99–138. Akadémiai Kiadó.