

MEGSZAKÍTOTT ÖSSZETEVŐK A MAGYARBAN KONSTRUKCIÓS NYELVTANI SZEMPONTBÓL

KÁLMÁN LÁSZLÓ ÉS RÁDAI GÁBOR

ELTE Elméleti Nyelvészet Szak

MTA Nyelvtudományi Intézete

Budapest I., Színház u. 5–9. Pf.: 19. 1250

E-mail: kalman@nytud.hu, radai@nytud.hu

0. Bevezetés

Ebben az írásban arra vállalkozunk, hogy az ún. *konstrukciós nyelvtan* (Construction Grammar) elméletét, eszközeit, lehetőségeit a magyar mondatban problémáin keresztül mutassuk be. Terjedelmi okokból mind a magyar nyelvtani jelenségek, mind a konstrukciós nyelvtan részleteit kissé lazán fogjuk kezelni, mert célunk elsősorban a megközelítés létjogosultságának és használhatóságának bemutatása, nem pedig a kézikönyv szintű ismertetés.

Először (az 1. részben) körvonalazzuk a magyar mondatannak azt a jelenségkörét, amely illusztrációul fog szolgálni. Arról van szó, hogy a magyarban a mondatrészek sorrendjét elsősorban a mondat kommunikációs szerepe, és ezen belül a különböző mondatrészek kommunikációs funkciója határozza meg, ami furcsa, a logikai szerkezettől különböző elrendezésekhez vezet. Ezenkívül a segédigés szerkezetek (pl. *be akar menni*) is sajátos szórendűek (az igekötő nem a saját igéje mellett helyezkedik el). A 2. részben röviden vázoljuk, hogy az ilyen ún. *megszakított* mondatösszetevők szokásos kezelési módjai milyen problémákat vetnek fel. Úgy látszik, hogy a létező elméletek matematikailag kezelhetetlen, *ad hoc* megoldásokat tartalmazó, ill. a tényleket nem megfelelően magyarázó kezelésmódokat tesznek csak lehetővé.

A 3. rész az írás középponti része: Ebben ismertetjük a konstrukciós nyelvtan fogalmát és az általunk javasolt eszközrendszert (3.1.), majd a 3.2. pontban tárgyaljuk, hogyan lehet ezek segítségével a nehézségeket okozó magyar jelenségeket leírni.

Már itt fontosnak tartjuk megemlíteni, hogy a konstrukciós nyelvtannak sem alapelveiben, sem eszközeiben nincs egyelőre egyetértés a különböző iskolák között. A konstrukciós nyelvtan alapművei (Hockett 1958, Fillmore és mtsai 1988) nem tartalmaznak részletesen kidolgozott mondatnani elméletet, a későbbi konkrét javaslatok (Fillmore–Kay 1993, Manaster-Ramer–Zadrozny 1992, Zwicky 1989) pedig egymástól lényegesen különböző elvárásokat és elveket fogalmaznak meg a konstrukciós nyelv-

tanokkal kapcsolatban. Az általunk ismertetett változat a saját elképzelésünket tükrözi, nem pedig ezeknek az előzményeknek a reprodukálása. Reméljük, nem minősül az elnevezéssel való visszaélésnek, ha mégis konstrukciós nyelvtannak nevezzük.

1. Néhány probléma

Köztudott, hogy a magyarban a szórend (és a főhangsúlyok helye) szorosan összefügg a mondatok ún. *aktuális* tagolásával. A mondatrészek közötti függőségi viszonyok, a mondat logikai szerkezete (az ún. predikátum-argumentum szerkezet) szinte elhanyagolható szerepet játszanak a mondatrészek sorrendjében. Azt elsősorban az határozza meg, hogy milyen kommunikatív szerepet tölt be a mondat a szövegben, s ezen belül milyen szerepet kapnak az egyes mondatrészek. Ennek illusztrálására tekintsük az alábbi egyszerű mondatokat:

- (1) a. 'Mari 'látta 'Jánost.
 b. 'Jánost 'látta 'Mari.
 c. 'Látta 'Jánost 'Mari! / 'Látta 'Mari 'Jánost!
 d. 'Látta Jánost Mari. / 'Látta Mari Jánost.
 e. 'Mari látta Jánost.
 f. 'Jánost látta Mari.
 g. 'Jánost 'Mari látta.
 h. 'Jánost 'látta Mari.

Az első mondat beszédértéke – tehát a szövegben betöltött kommunikatív szerepe – szerint semleges közlés, vagyis egyszerűen új információt közöl, míg a mondatrészek szerepe a következőképpen alakul: a mondat topikja – tehát akiről mondunk valamit – az alany, vagyis Mari, míg a komment – vagyis amit mondunk róla, az új információ – szerepét a mondat többi része, látta Jánost, tölti be. A következő két példa, a (b) és a (c) beszédértéke szintén semleges közlés, de megváltozik a mondat topik-komment szerkezete. Míg az előbbi esetben a mondat tárgya áll topikban, vagyis Jánosról állítunk valamit – mégpedig azt, hogy Mari látta őt –, addig az utóbbi mondatban nem szerepel topik, és az egész kijelentés tölti be a komment szerepét. A felkiáltójelel is utal rá, hogy itt egy, a beszélgetés kontextusából kilógó, váratlan hírrrel van dolgunk. A további példák esetében a beszédérték már nem semleges, ezek a mondatok visszavágásként szolgálnak valamely őket megelőző kijelentésre. Ezen belül a (d) mondat két variánsa – amelyek viszont egyenértékűek – nyomatékosítást fejez ki, vagyis egy olyan kijelentésre reagál, miszerint nem látta Mari Jánost, és azt állítja, hogy ez nem igaz, vagyis hogy Jánost igenis látta Mari. Ezzel szemben az (e) és (f) példákban egy újabb jelenséggel, a *fókusszal* találkozunk, ahol is az első esetben Mari, míg a másodikban János a fókuszált összetevő, és ezek a mondatok egy olyan kontextusban szolgálhatnak válaszként, ahol nem az egész tényállást, csak annak valamely momentumát

kívánja a beszélő megerősíteni. Így az (e) példa akkor szolgálhat adekvát válaszul, ha valaki kijelenti, hogy valaki látta Jánost, és vagy nem tudja, hogy ki, vagy mást nevez meg, és hasonlóképpen működik az (f) mondat is. Az utolsó két példa a *kontrasztív topik* jelenséget illusztrálja: ezekben a mondatokban a *János* van ebben a pozícióban, ami azt jelenti, hogy a beszélő őt kívánja szembeállítani más személyekkel. A kontrasztív topiknak az a szerepe, hogy a mondat nem kimerítő voltát emelje ki. Például a (g) mondattal reagálhatunk arra a kijelentésre, hogy Jánost és másokat Mari látott, és hangsúlyozhattuk, hogy Jánost ugyan Mari látta, de ez nem a teljes igazság: másokat esetleg nem. Hasonlóképpen válasz lehet a (g) mondat arra a kérdésre, hogy *Ki látta a fiúkat?*, ha csak arra tudunk, vagy csak arra akarunk válaszolni, hogy a fiúk közül Jánost ki látta. A (h) mondat hasonlóképpen azt állítja, hogy Jánost ugyan látta Mari, de másokat esetleg nem, és válasz lehet arra a kérdésre is, hogy *Kit látott Mari?*, ha csak egyet tudunk vagy akarunk megnevezni a kért személyek közül.

A magyar mondatoknak az a sajátossága, hogy szórendjük kommunikációs szerepüktől függ, eleve komoly kihívást jelent a hagyományos mondattani elméleteknek, amelyek főleg a logikai függőségi viszonyok alapján igyekeznek megfogalmazni a szórendi szabályokat, hiszen azokban a nyelvekben, amelyek leírására eredetileg kidolgozták őket, valóban elsősorban ez határozza meg az összetevők felszíni szerkezetét.¹ Később visszatérünk még arra, hogy ez milyen nehézségeket okoz.

Az aktuális tagolás hatása a magyar szórendre abban is megnyilvánul, hogy olyan mondatrészek, amelyek egymással állnak a legszorosabb logikai kapcsolatban, bizonyos esetekben egymástól távol, elszakítva helyezkednek el a mondatban. Ezek a mondatban egy szintaktikai egységet, összetevőt alkotnak (legalábbis az összetevők hagyományos értelmében). Különleges szórendi sajátosságaik miatt *megszakított összetevőknek* kellene őket tekintenünk. Vessünk újra egy pillantást az (1g) példára, ahol megszakított összetevőknek kell tekintenünk a mondat állományát, vagyis a *[[Jánost]...[látta]]* szerkezetet, amely egyrészt logikailag összetevőt alkot – ti. egy predikátumot, amelyet Mariról állítunk –, másrészt elemei a semleges mondatokban (ld: (1a–b)) szomszédosak.

A megszakított összetevők mindenképpen problémákat okoznak a generatív nyelvészetben megszokott mondattani elméleteknek, amelyekben (akárcsak a hagyományos nyelvtanban) az alapműveletek mondatrészek *egymásba ágyazása és egymás mellé helyezése*. (Megszakított összetevők esetén olyan mondatrészek vannak egymás mellé helyezve, amelyek közvetlenül nem ugyanabba a mondatrészbe vannak beágyazva.) Van azonban egy jelenség a magyarban, ami végképp óriási gondot jelent a legtöbb mondattani elmélet számára: az igekötők „*elszakadása*” az igétől, amit illusztrál

¹Ezt a kifejezést itt intuitív, preteoretikus értelmében használjuk.

a következő példa, ahol az igekötőt *hátravetjük* a fókusz után, és ahol az aktuális mondattagolás hatásaként lép fel ez a jelenség.

(2) *'Jánost a 'vonat ütötte el.*

De tekinthetjük az alábbi példát is, ahol a közlésérték semleges, de a segédige *beférkőzése* az igekötő és infinitívusza közé mégis elszakítja a két összetevőt egymástól.

(3) *'Jánost 'el fogja 'ütni a 'vonat.*

Az egyik ok, amiért az igekötő elválása nagyobb gondot okoz a mondatrészek „vándorlásánál”, hogy gyakran szótárilag összetartozó, nem áttetsző jelentésű kapcsolatokat szakít ketté (pl.: *be fog rúgni*); vagyis olyan szerkezeteket, amelyeknek jelentése nem kompozicionális, részeinek jelentéséből nem kiszámítható, és amelyeknek éppen ezért egy egységet kell alkotniuk a lexikonban (a szó technikai értelmében). Igaz, nemcsak igekötős igékkel, hanem szószerkezetekkel is előfordul, hogy nem áttetsző a jelentésük, mégis megszakított összetevőt alkotnak (pl. *[[húzni] fogja [a lóbört]]*), de az igekötős igék sok szempontból az összetett szavakhoz hasonlóak, kettészakadásuk ezért különösen meglepő.

2. Néhány megoldási lehetőség

Az 1. pontban említett problémák valamennyien a mondattani szerkezet két- vagy többarcúsága mellett tanúskodnak: a logikai szerkezet különválík a lineáris szórendtől és az aktuális tagolás szerkezetétől; az aktuális tagolás közelebbi kapcsolatban van a lineáris szórenddel, mint a logikai szerkezet, de (a segédigék beférkőzésénél) mindhárom különválík. Kézenfekvő lenne tehát, hogy az adatokat úgy magyarázzuk, hogy többféle szerkezetet és ezek között közvetítő szabályokat (például mozgó szabályokat) feltételezzünk.

Ezeket a megoldásokat azonban a következőkben nem fogjuk vizsgálni, egyrészt azért, mert a több elemzési szinttel dolgozó nyelvtanok általában egy sor *ad hoc* elemet tartalmaznak, másrészt pedig, mert matematikailag nehezen kezelhetőek, s ezért csak végső eszközként használhatnánk őket, ha a konzervatívabb megoldások csődöt mondanának. Másképpen kifejezve: ha többszintű elemzéssel – pl. *transzformációs generatív* mondattani keretben, vagy a *lexikai-funkciós nyelvtanban* (az ún. LFG-ben) – reprodukálni tudjuk az általunk javasolt konzervatívabb elemzést, az az égvilágon semmit sem jelent.

Nem kívánjuk ugyanakkor tagadni azt az általánosítást, hogy a vizsgált jelenségek különböző (logikai, aktuális tagolásbeli és esetleg bizonyos elemekhez, pl. a segédigékhez kapcsolódó sajátos) szerkesztési elvek összjátékából fakadnak. A mondattannak ezzel a sajátosságával viszont azoknak az elméleteknek van gondjuk, amelyek nem támaszkodnak egynél több szintre a mondattani elemzésnél. Ezek az elméletek

minden mondatnak egyféle, a mondat felszíni felépítését magyarázó szerkezetet tulajdonítanak, és ezért a mondat szerkezet felépítéséért felelős szabályokat is egységesnek szokták tekinteni (pl. az ún. *frázisstruktúra-nyelvtanokban* egyforma *újraíró* szabályokat feltételeznek, a *kategoriális* nyelvtanokban pedig szomszédos elemeken végrehajtott függvénytáblázatokat). Ezekben az elméletekben a mondatok jelentésével is a mondat szerkezet felépítésével párhuzamosan kell elszámolni, úgy, hogy minden összeépítési szabály vagy lépés jelentés-összeépítéssel jár.

Az egyszintű elemzésekben a távoli függőségeket lényegében egyféleképpen lehet kezelni: olyan sajátos mondattani kategóriák feltételezésével, amelyek összetettek, és valóságos „memóriaként” tárolják az információt egymástól távol eső mondatrészek összetartozásáról:

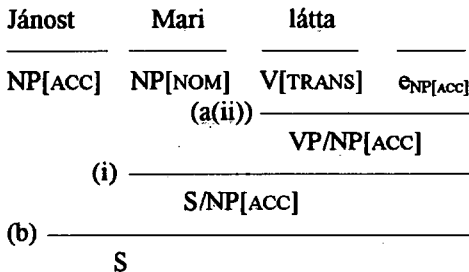
(4) *Újraíró szabályok:*

- (i) $S[\alpha] \rightarrow NP[NOM] + VP[\alpha]$
- (ii) $VP[\alpha] \rightarrow V[TRANS] [\alpha] + NP[ACC]$

Metaszabályok:

- (a) Hiányos kifejezés: $X \rightarrow Y + Z \Rightarrow X/Z \rightarrow Y + e_z$
- (b) Hiány megszűnése: $Z \rightarrow Y + Z/Y$

Egy levezetés:



Ebben az elemzésben, amely nagyjából egy GPSG-beli (*általánosított frázisstruktúra-nyelvtani*) elemzésnek felel meg, a $VP/NP[ACC]$ hiányos igei csoportot jelöl (egy tárgy esetű főnévi csoport, $NP[ACC]$ hiányzik belőle). Az $S/NP[ACC]$ hasonlóan hiányos mondatot jelöl; a mondat (S) a kontrasztív topikban levő tárgyból ($NP[ACC]$) és a hiányos mondatból tevődik össze. A távoli függőség így válik „közelivé”.

A következő elemzés egy *kategoriális* nyelvtani levezetés lehetne:

(5) Jánost	Mari	látta
NP[ACC]	NP[NOM]	(S/NP[NOM])/NP[ACC]
TR		C
S/(S/NP[ACC])		(S/NP[ACC])/NP[NOM]
	BA	
	S/NP[ACC]	
A		
S		

Itt a *látta* kategóriája a tárgyas igéké: ha jobbról kap egy tárgyat (/NP[ACC]), akkor olyan kategóriát ad (S/NP[NOM]), amely balról vár egy alanyesetű főnévi csoportot, és mondatot eredményez. A C nevű művelet megcseréli az argumentumhelyeket (így balról az alanyt, majd jobbról a tárgyat váró ige kategóriáját kapjuk: (S/NP[ACC])/NP[NOM]). A BA („függvényalkalmazás balra”) révén az alany helye kitöltődik, s a *Mari látta* olyan kategóriát kap, mely jobbról vár egy tárgyat, hogy mondatot kapjunk (S/NP[ACC]). A TR („topic raising”) nevű művelet a tárgyesetű főnévi csoporthoz olyan kategóriát rendel, amely egy tárgyhiányos mondatot (S/NP[ACC]-ot) vár jobbról. Az A („függvényalkalmazás jobbra”) segítségével kapjuk meg végül az eredményt, hogy az egész kifejezés mondat.

Mindkét elemzés azt a nyilvánvaló tényt fogalmazza meg formálisan, hogy a kontrasztív topiknak a mondat elejére kell kerülnie, még ha a logikai szerkezetben valahová a végére tartozna is. Ugyanakkor olyan elem is állhat kontrasztív topikban, amely egyáltalán nem a mondat végétől „kerül oda”, sőt, esetleg nélküle a mondat nem is lenne hiányos. Ezeket sajnos nem lehet a fenti esettel egységes módon kezelni.

További csúnya hibája az ilyen típusú megközelítéseknek, hogy jelentősen túlgenerálnak, mégpedig több szinten is. Egyrészt az ilyen eszközök megengedése, amelyek segítségével tárolható valamely mondatrész hiányáról szóló információ, ami az egyre nagyobb szerkezetek összerakásánál felöröklődik, és később bármely ponton megszüntethető, ahogyan ez a kategoriális nyelvtanban történik, ún. permutációzárt-sághoz vezet, vagyis a nyelvten a szavak bármely sorrendjét grammatikailag helyes mondatnak fogadja el, ami még az olyan köztudottan szabad szórendű nyelvek esetén is túlzás, mint amilyen a magyar. Csak további megszorításokkal lehet ennek elejét venni, ami persze jelentősen bonyolítja a nyelvten. Ami a GPSG-féle rendszereket illeti, ahol jegyek formájában kódolják a hiányosságot (s ezek a jegyek öröklődnek), ezekben a nem terminális szimbólumok (vagyis mondattani kategóriák) elvben végtelen számúak, ami matematikailag szinte kezelhetetlen (úgynevezett kontextusfüggő) nyelvten eredményez.

Másrészt szemantikailag is túlgenerál a nyelvten, amennyiben – éppen az előállított elemzések nagy száma miatt – a grammatika egyébként valóban a nyelvhez tartozó mondatok esetében olyan elemzéseket is produkál, amelyek szemantikai

megfelelője bizonyosan nem jelentése a mondatnak. A fenti példánál maradva, olyan mondatrész is kaphat kontrasztívtopik-értelmezést, amely teljesen független okokból került mondatvégi helyzetből a mondat elejére.

A legnagyobb probléma azonban a mondat szerkezet és a jelentés párhuzamos előállításával van. Például – mint arra már fentebb rámutattunk – az igekötő elválásakor az igekötő és az ige jelentését külön kellene tartanunk (mert két külön összetevőként kell működniük a mondat szerkezetben, ha elválnak), akkor is, amikor ez lehetetlen, mert az igekötő és az ige nem transzparens jelentésű szótári egységes alkotnak.

3. Konstrukciós nyelvtani megoldások

3.1. A konstrukciós nyelvtan fogalma

3.1.1. Vezérelvek

A *konstrukciós nyelvtannak* (Construction Grammar) két fontos jellegzetessége van:

- (i) *Egyszintű*: a kifejezéseket felépítő (a kifejezésekben megtestesülő) *konstrukciók* közvetlenül jelentésbeli és formai sajátosságok összekapcsolódását fejezik ki, közbeeső szintek nélkül.
- (ii) A szokásos egyszintű elméletekkel szemben a kifejezések szerkezetét *nem szigorú hierarchiaként* képzelem el. Vagyis a kifejezést nem kell feltétlenül közvetlen összetevőik egymásutánjaként (*konkatenációjaként*) felfognunk. A kifejezések alakja (és jelentése) egy vagy több konstrukció formai (és jelentésbeli) követelményeinek együttes kielégítéséből fakad, s ezek a konstrukciók nemcsak valamilyen hierarchiában egymás mellé rendelt elemek konkatenációját írhatják elő.

A későbbiekben látni fogjuk, hogy milyen következményekkel járnak ezek a módszertani elvek.

3.1.2. Jegy-érték struktúrák

Az általunk használt konstrukciós nyelvtani formalizmusban minden konstrukciót *jegy-érték struktúra* (attribute-value structure, röviden *AVS*) formájában fogunk felírni.² Az ilyen struktúrákban jegy-érték párok vannak felsorolva, úgy, hogy minden jegyhez csak egy érték tartozhat, és az értékek vagy atomiak, vagy maguk is AVS-ek. Minden AVS objektumok egy halmazát írja le – denotálja –, mégpedig azon adott

²Az AVS-ek formális definiálásától eltérően, mivel ennek számos változata megtalálható a szakirodalomban: ld. Johnson (1988).

típusú objektumok halmazát, amelyeknél a megfelelő jegyek az AVS-ben szereplő értéket veszik fel. Lássunk egy egyszerű AVS-t:

[CAT NP]

Ez a leírás megfelelhet azoknak a szintaktikai szerkezeteknek, amelyeknek kategóriája főnévi csoport (a CAT, vagyis a „kategória”-jegy értéke az atomi NP). Ezen a halmazon belül a tárgyesetű főnévi csoportok a

[CAT NP
CASE ACC]

struktúrával jellemezhetők. Hasonlóképpen értelmezhető, ha egy jegy értéke maga is AVS:

[SUBJ [CAT NP
CASE NOM
PERS 3]]
PRED [CAT V]]

Ez az AVS olyan struktúrák leírására szolgálhat, melyeknek alanya alanyesetben álló, harmadik személyű főnévi csoport, állítmánya pedig ige. Vagyis az összetett AVS-ek összetett jegyek értékeit adják meg, mint amilyen a fenti példában az „alany esete”, a SUBJ CASE.

Az A és B AVS-ek unifikációján ($A \sqcup B$) olyan AVS-t értünk, amely az A által jelölt objektumhalmaz és a B által jelölt objektumhalmaz metszetét jelöli. Például az fent említett két első AVS unifikációja azonos a másodikkal:

[CAT NP
CASE ACC]

Ez megfelel annak, hogy a második AVS által meghatározott halmaz részhalmaza az első AVS-t kielégítő halmaznak, vagyis kettejük metszete azonos a második AVS által jelölt halmazzal.

Az AVS-eknél általánosan használatos további eszköz a *közös struktúra* (structure sharing), ami azt jelenti, hogy bizonyos jegyértékek nemcsak ugyanazt az információt hordozzák, de ténylegesen azonosak. Ezt bekeretezett indexekkel jelölik:

[SUBJ [CAT NP
CASE NOM
PERS 1]]
PRED [CAT V
INFLPERS 1]]

Ez az AVS olyan objektumokat (feltehetően mondatokat) ír le, amelyeknek alanya alanyesetű főnévi csoport, és állítmánya olyan ige, amelynek személyragja megegyezik az alany személyével.

Mivel minden konstrukció jelentés és forma összekapcsolódását fejezi ki, az általunk használt AVS-ek

$$\left[\begin{array}{l} \text{MEANING ...} \\ \text{FORM ...} \end{array} \right]$$

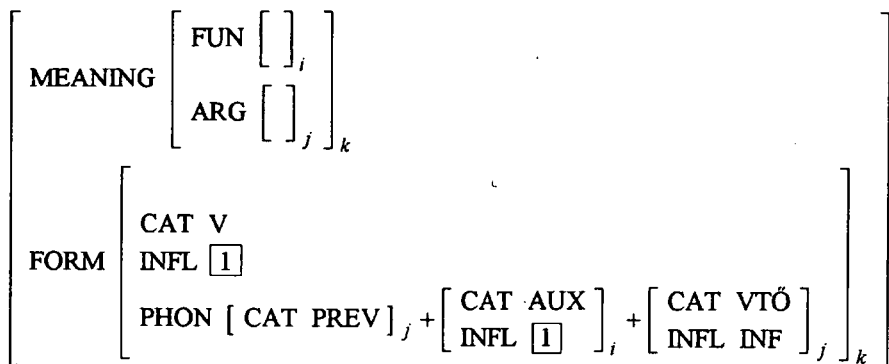
alakúak lesznek, ahol értelemszerűen a MEANING jegy értéke a konstrukció jelentésének valamilyen leírását tartalmazza, míg a FORM jegy értéke azon szintaktikai szerkezetek halmazát denotálja, amelyek ezt a jelentést hordozzák.³

Még további két sajátos eszközt fogunk ezeken kívül használni. Az egyik a *koindexálás*, amit arra fogunk használni, hogy bizonyos jelentésbeli és formai egységeket megfeleltessünk egymásnak:

$$\left[\begin{array}{l} \text{MEANING} \left[\begin{array}{l} \text{FUN} \left[\begin{array}{l} \\ \end{array} \right]_i \\ \text{ARG} \left[\begin{array}{l} \\ \end{array} \right]_j \end{array} \right]_k \\ \text{FORM} \left[\begin{array}{l} \text{CAT VP} \\ \text{PHON} \left[\text{CAT V} \right]_i + \left[\begin{array}{l} \text{CAT NP} \\ \text{CASE ACC} \end{array} \right]_j \end{array} \right]_k \end{array} \right]$$

Ez az AVS olyan jelentés-forma összekapcsolódást ír le, amelyben a jelentés egy tetszőleges (*i* indexű) ige jelentéséből és egy másik tetszőleges (*j* indexű) tárgyesetű főnévi csoport jelentéséből tevődik össze úgy, hogy az első a FUN („függvény”) jegy értéke, míg a második az ARG („argumentum”) jegy értéke. A felszíni alakot a FORM PHON összetett jegy értéke írja le, a jelen esetben az írja elő, hogy az *i* indexű elemet közvetlenül kövesse a *j* indexű elem. A „+” jel tehát konkatenációt jelöl. A koindexálást azért nem helyettesítheti a közös struktúra, mert olyan szavak, egységek is viselhetnek azonos indexet, amelyek egymással nem azonosak (összeegyeztethetetlen információt tartalmaznak):

³A továbbiakban minket elsősorban a mondatok felszíni sorrendje fog érdekelni, ezért a MEANING jegy értéke csak valamilyen sematikus struktúra lesz.



Ez az AVS olyan szerkezetet ír le, amelynek jelentésében a FUN jegy értéke az i indexű segédige (az INFL jegy azt fejezi ki, hogy milyen alakú az ige, milyen *inflexiót* visel), az ARG értéke pedig a j indexű igekötős ige. A formát illetően azonban a segédige az igekötő és az igető között helyezkedik el, amelyek mindketten a j indexet viselik a PHON értéken belül.⁴

A másik sajátos eszközre a fent már látott „+” jelölés a példa: a formai jellemzéseken belül használni fogunk egy formális nyelvet, amelynek segítségével a felszíni formára vonatkozó formai megszorításokat ábrázoljuk. A „+” jelen kívül használni fogjuk még a halmazelméletből ismerős „{...}” jelölést is, amely a kapcsos zárójelben álló elemek tetszőleges sorrendjére utal, úgy, hogy ha egy ilyen felsoroláson belül hasonló halmazjelölésű felsorolás állna, akkor a belső halmaz kapcsos zárójelei eltűnnek, a két lista elemei „összekeveredhetnek”. A más listák elemeivel „nem összekeveredő”, de tetszőleges szórendű részek listáját kettős kapcsos zárójelbe tesszük: „{{...}}”. A választhatóan jelen levő részeket zárójelbe tesszük. Végül használni fogjuk a „&” jelet is két olyan megszorítás összekapcsolására, amelyeket a felszíni alaknak egyszerre kell teljesítenie. A következő alponthan tömören összefoglaljuk a PHON jegy értékében használt alnyelvet. Előrebocsátjuk, hogy a felszíni alakok leírásában sem a középpontozást (pont, vessző, szóköz), sem a hangsúlyozást nem fogjuk ábrázolni.

3.1.3. Stringleírások

Ennek az alponthanak a megértése nem feltétlenül szükséges a további elemzések megértéséhez. A formális nyelvi modellekben nem járatos olvasó átugorhatja.

Mint említettük, a konstrukcióknak megfelelő felszíni stringek halmazát a PHON jegy értéke jelöli az egyes konstrukciókban. A PHON jegy értékében egy sajátos nyelvet használunk, amelyben nyelvtani (kategorialis) információ és közvetlen fonológiai információ keveredik. Ezt a nyelvet azonban le tudjuk fordítani egy tisztán

⁴Erre a mechanizmusra szükségünk is lesz az igekötős igék szétszakadásának leírásánál, ha a fenti AVS szó szerint nem fog is előfordulni.

stringleíró nyelvre, amelynek minden kifejezése stringek egy *reguláris* halmazát jelöli. (Minden reguláris stringhalmaz azonos valamely *véges állapotú automata* által elfogadott stringhalmazzal, és viszont. A reguláris halmazok zártak a metszet műveletére, ezért a PHON jegy értékeinek unifikációját a nekik megfelelő véges automaták metszeteként értelmezhetjük.)

A PHON jegy értékeiben szereplő kifejezésekből a stringleíró kifejezéseit úgy kapjuk, hogy a bennük szereplő AVS-ek helyett a bennük szereplő PHON jegy értékét vesszük (ha nem szerepel bennük a PHON jegy, akkor a [] speciális leírást), egyetlen megszorítással: ha a bennük szereplő PHON jegy értéke $\{\alpha_1, \dots, \alpha_n\}$ alakú, és az AVS maga is egy ilyen felsorolás egyik eleme, akkor az α -kat körülvevő kapcsos zárójelet elhagyjuk, mert ez felel meg a $\{\cdot\}$ kívánt értelmezésének. (Mellesleg ez történik akkor is, ha egy $\{\dots\}$ alakú PHON érték egyik elemébe egy másik $\{\dots\}$ alakút kell beépítenünk.)

A kapott stringleíró kifejezések szintaxisa és szemantikája a következő. A stringleíró kifejezések nyelve

$$\mathcal{L} =_{\text{def}} \langle \text{LC}, \text{Con}, \text{Exp} \rangle,$$

ahol LC a logikai konstansok halmaza:

$$\text{LC} =_{\text{def}} \{ [], +, \{, \}, \{ \{, \} \}, \&, (,) \},$$

Con a nem logikai konstansok halmaza:

$$\text{Con} =_{\text{def}} V^* \setminus \{\varepsilon\},$$

(itt V^* az ábécé fölötti összes véges stringek halmaza, ε pedig az üres string; az egyszerűség kedvéért a nem üres stringeket bevesszük a tárgynyelvbe konstansként). Exp a jól képzett kifejezések halmaza. Exp a legkisebb halmaz, amely eleget tesz az alábbi (i–vi) szintaktikai szabályoknak. Az (i'–vi')-ben megadjuk az adott szintaktikai szabály szerint képzett kifejezések szemantikai értékét. ($\cdot \mid \cdot$). A modell kizárólag a V^* halmazból áll, a $\mid \cdot \mid$ függvény minden jól képzett kifejezéshez V^* egy részhalmazát rendeli.⁵

A jól képzett kifejezéseket két részre osztjuk, a $\{\dots\}$ alakú kifejezéseket az Exp. halmazba, a többi az Exp_0 halmazba soroljuk:

$$\text{Exp} =_{\text{def}} \text{Exp}_s \cup \text{Exp}_0,$$

$$\text{Exp}_s \cap \text{Exp}_0 = \emptyset.$$

⁵Mivel a konstansok halmaza V^* részhalmaza, ezért a szintaktikai algebra nem szabad algebra. Könnyű azonban ellenőrizni, hogy az interpretáló függvény megválasztása miatt – például a szintaktikai konkatenációnak a szemantikai algebrában is konkatenáció felel meg (Id. (iii) ill. (iii') – a két algebra mégis homomorf lesz.

Erre azért van szükség, hogy a kapcsos zárójeleken belül közvetlenül a szintaxis szerint se szerepelhessenek kapcsos zárójeles kifejezések.

- (i) $\alpha \in \text{Con} \Rightarrow \alpha \in \text{Exp}_0$.
Minden string egyben jól formált kifejezés.
- (i') $|\alpha| =_{\text{def}} \{\alpha\}$.
Minden string a belőle álló egyelemű halmazt jelöli.
- (ii) $[\] \in \text{Exp}_0$.
A $[\]$ jel jól képzett kifejezés.
- (ii') $|\ [\] | =_{\text{def}} V^*$.
A $[\]$ tetszőleges stringet jelölhet (az üres stringet is).
- (iii) $\alpha_1, \dots, \alpha_n \in \text{Exp} \Rightarrow \alpha_1 + \dots + \alpha_n \in \text{Exp}_0$.
Kifejezések konkatenációja.
- (iii') $|\alpha_1 + \dots + \alpha_n| =_{\text{def}} \{s_1 \dots s_n \in V^* : s_1 \in |\alpha_1| \wedge \dots \wedge s_n \in |\alpha_n|\}$
A konkatenáció az egyes halmazokból vett stringek konkatenációinak halmazát jelöli.
- (iv) $\alpha_1, \dots, \alpha_n \in \text{Exp}_0 \Rightarrow \{\alpha_1, \dots, \alpha_n\} \in \text{Exp}_s, \{\{\alpha_1, \dots, \alpha_n\}\} \in \text{Exp}_0$.
A kétféle kapcsos zárójeles kifejezések jól formáltak, ha az elemeik nem maguk is kapcsos zárójelesek.
- (iv') $|\{\alpha_1, \dots, \alpha_n\}| =_{\text{def}} |\{\{\alpha_1, \dots, \alpha_n\}\}| =_{\text{def}} \{s_1 \dots s_n \in V^* : \bigwedge_{1 \leq i \leq n} \bigvee_{1 \leq j \leq n} (s_i \in |\alpha_j|) \wedge \bigwedge_{1 \leq i \leq n} \bigvee_{1 \leq j \leq n} (s_i \in |\alpha_j|)\}$.
Minkét kapcsos zárójeles kifejezés értelmezése: a bennük szereplő stringhalmazok egy-egy elemének tetszőleges sorrendben való konkatenációja.
- (v) $\alpha_1, \dots, \alpha_n \in \text{Exp} \Rightarrow \alpha_1 \& \dots \& \alpha_n \in \text{Exp}_0$.
Stringléírások mint követelmények együttes érvényesítése.
- (v') $|\alpha_1 \& \dots \& \alpha_n| =_{\text{def}} \bigcap_{1 \leq i \leq n} |\alpha_i|$.
Az alapértelmezésük a megfelelő stringhalmazok metszete.
- (vi) $\alpha \in \text{Exp} \Rightarrow (\alpha) \in \text{Exp}_0$.
Választhatóan megjelenő string.
- (vi') $|(\alpha)| =_{\text{def}} |\alpha| \cup \{\varepsilon\}$.
Értelmezése: vagy a stringhalmaz valamelyik eleme, vagy az üres string.

3.1.4. Műveletek konstrukciókkal

A konstrukciós nyelvtan alapján való *elemzés* (parsing) a nyelvtan természeténél fogva *alakfelismerési* feladat, hiszen azt kell megállapítani, hogy milyen konstrukciók testesülnek meg egy adott kifejezésben. Ez az alakfelismerési feladat a vizuális alakfelismeréshez hasonló, megoldására nem vállalkozunk. Csak arra teszünk kísérletet, hogy bizonyos rendelkezésünkre álló konstrukciók alapján mondatokat *generáljunk*, mégpedig úgy, hogy egy ismert jelentéshez próbálunk formát találni. Azzal sem fogunk itt foglalkozni, hogy a konstrukciók tárháza (a voltaképpeni nyelvtan) hogyan épül fel. Csak azt a két műveletet fogjuk tárgyalni, amelyek segítségével konstrukciók összeépíthetőek.

Mint már említettük, a hagyományos nyelvtanokkal szemben az összeépítés nem korlátozódik egymásnak mellérendelt szerkezetek konkatenációjára (és egyben a nagyobb szerkezetbe való hierarchikus beépítésre, alárendelésre). Ehelyett a használt két művelet a következő lesz:

- (a) Két konstrukció *egyidejű* megvalósítása. Ennek a két megfelelő AVS unifikációja felel meg.
- (b) Egy konstrukció beépítése egy másikba. Ez is unifikációs lépésekből áll, lényege, hogy a beépített konstrukció MEANING értékét a nagyobb konstrukció MEANING értékén belül egy bizonyos indexű érték(ek)kel unifikáljuk, és a FORM értékét a nagyobb konstrukció FORM értékén belül az ugyanazzal az indexszel megjelölt érték(ek)kel unifikáljuk.

A következő alpontban több példát is látunk majd ezekre a műveletekre. Csak annyit jegyzünk meg meg, hogy bizonyos különleges esetekben a FORM értékekben megfogalmazott formai követelmények egyesítése nem egyszerű unifikáció. Ez a (b) típusú beépítéseknél fordulhat elő. Ésszerűnek látszik ugyanis feltételezni, hogy ha formai követelmények konfliktusba kerülnek egymással, akkor a nagyobb szerkezet követelményei elsőbbséget élveznek. Ez az ún. *Felülről Lefelé Elsőbbség* (*Top-Down Priority, TDP*) elve. Például az angolban, ahol a tárgynak általában közvetlenül az ige után kell állnia, néhány mondatszintű konstrukció (kiegészítendő kérdés, topikalizáció) ezt a követelményt felül tudja bírálni.

Még egy megjegyzést kell tennünk a jelentésből való mondatgenerálással kapcsolatban. Látni fogjuk, hogy ahhoz, hogy a kívánt felszíni formák álljanak elő, elég speciális jelentésábrázolásból kell kiindulnunk. Például a magyar esetében már a jelentésábrázolásba bele kell írunk, hogy melyik főnévi csoport tölti be a topik szerepét. Ha más nyelven akarnánk mondatokat előállítani, akkor erre esetleg nem lenne szükség, illetve más részletekre lehet szükség a jelentésábrázolásban, amelyekre a magyar esetében nem hivatkozunk. Első látásra különös lehet, hogy a különböző nyelv-

vekben a kiinduló jelentésábrázolások nem azonosak. Úgy tűnhet, hogy ettől kissé önkényes lesz a konstrukciókon alapuló generálás, hiszen mindig olyan jelentésábrázolást tételezhetünk fel, amelyből a kívánt felszíni kifejezés „kijön”. Másrészt igenis indokolt, hiszen a különböző nyelvek különböző jelentéseket fejeznek ki, legalábbis különböznek a *kötelezően kifejezendő* jelentésvonások. Ezért kénytelenek vagyunk feltételezni, hogy a jelentésábrázolásoknak legalábbis van egy olyan szintje, amely valóban nyelvenként változik, és ezekből kell kiindulnunk a generálásnál.

3.2. Magyar szerkezetek előállítás konstrukciókból

A magyar nyelvtan számos vonását figyelmen kívül fogjuk hagyni az alábbi elemzésekben, mivel már az általunk vizsgált problémák is elég bonyolult ábrázolásokat tesznek szükségessé. Ennek megfelelően az 1. pontban idézett példákban szereplő szótári egységekhez tartozó konstrukciókat is leegyszerűsített formában ábrázoljuk. (Felhívjuk a figyelmet arra, hogy a szótári egységek, mint alapvető jelentés-forma társítások, szintén konstrukciók!)

3.2.1. Szótári egységek

A következőkben megadjuk az (1) példában szereplő mondatok leírásához szükséges szótári egységeket.

$$(6) \quad \left[\begin{array}{l} \text{MEANING John}_{14} \\ \text{FORM} \left[\begin{array}{l} \text{CAT NP} \\ \text{CASE ACC} \\ \text{PHON } \textit{Já nost} \end{array} \right]_{14} \end{array} \right]$$

Mint látható, a *Jánost* szót egy szótári egységnek tételezzük, ami már tartalmazza a tárgyeset ragját, illetve az ennek megfelelő jegyet. Ennek oka, hogy jelen cikk témája nem a szótár felépítése, illetve a morfológia. Hasonlóképpen nem kívánunk a jelentéskomponens leírásában mélyebben belemenni, ezért választottuk jelentésként a szavak angol megfelelőit.

Hasonlóan épül föl a *Marinak* megfelelő szótári egység is:

$$(7) \quad \left[\begin{array}{l} \text{MEANING Mary}_{15} \\ \text{FORM} \left[\begin{array}{l} \text{CAT NP} \\ \text{CASE NOM} \\ \text{PHON } \textit{Mari} \end{array} \right]_{15} \end{array} \right]$$

Szükségünk lesz még a *látta* ragozott igét megjelenítő analízatlan és leegyszerűsített szótári egységre:

$$(8) \quad \left[\begin{array}{l} \text{MEANING } \textit{saw}_{16} \\ \text{FORM} \left[\begin{array}{l} \text{CAT V} \\ \text{INFL } + \\ \text{PHON} \left[\begin{array}{l} \text{CAR } + \\ \text{INFL } + \\ \text{PHON } \textit{lá tta} \end{array} \right]_{16} \end{array} \right] \end{array} \right]_{16}$$

Mint fentebb jeleztük, nem fogjuk tárgyalni azokat a morfológiai folyamatokat, amelyek létrehozzák a ragozott alakokat (látta: múlt idő, egyes szám harmadik személy, tárgyas ragozás; látni: infinitívusz). Sőt, a ragozott alakokban az INFL értékét sem fogjuk részletesen feltüntetni, ez utóbbi jegyet kétértékűnek tételezzük; a „+” érték jelen esetben azt jelenti, hogy ragozott alakkal van dolgunk. Az igen PHON értékén található CAR jegy – amely megint csak kétértékű – nagyjából azt jelzi, hogy szerepelhet-e az adott elem a komment elején (lehet-e ún. *igevivő*: a *látta* ragozott ige lehet, míg például a *fogja* segédige nem). Ezt a későbbiekben részletesen elmagyarázzuk, és ott látni fogunk konkrét példát a jegy felhasználására.

További két szerkezet szótári alakját is megadjuk, mert a továbbiakban szükségünk lesz rá. Az a *vonat* főnévi csoport szintén nem szótári egység ugyan, mégis egyetlen konstrukcióként kezeljük, mert belső felépítése nem fontos a számunkra:

$$(9) \quad \left[\begin{array}{l} \text{MEANING } \textit{the train}_{17} \\ \text{FORM} \left[\begin{array}{l} \text{CAT NP} \\ \text{CASE NOM} \\ \text{PHON } \textit{a vonat} \end{array} \right]_{17} \end{array} \right]$$

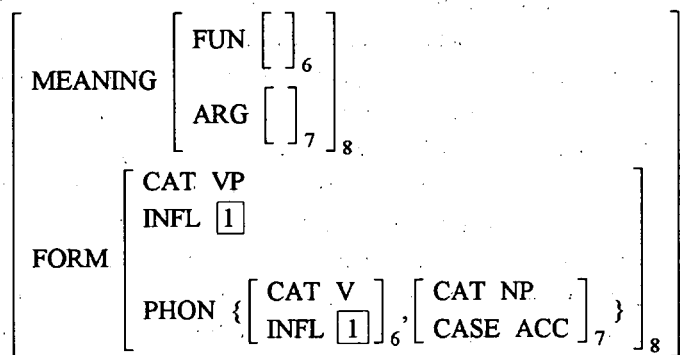
Az *elütni* (valamint *ütni* (...) *el* és *el ...ütni*) igekötős igeinek azonban fontos a belső szerkezete, noha szótári egység.

$$(10) \quad \left[\begin{array}{l} \text{MEANING } \textit{hit}_{18} \\ \text{FORM} \left[\begin{array}{l} \text{CAT V} \\ \text{INFL } - \\ \text{PHON} \left\{ \left[\begin{array}{l} \text{CAT PREV} \\ \text{CAR } + \\ \text{PHON } \textit{el} \end{array} \right]_{18}, \left[\begin{array}{l} \text{CAT V} \\ \text{PHON } \textit{ütni} \end{array} \right]_{18} \right\} \end{array} \right]_{18} \end{array} \right]$$

Látható, hogy az *elütni* szótári reprezentációja két elemből áll, ez biztosítja, hogy adott konstrukciókban az igekötő valóban „elszakadhasson” az igétől. Feltételeztük azt is, hogy az *elütni* ige nem áttetsző jelentésű, tehát mindkét eleme ugyanazzal a jelentéssel

mondattal állunk-e szemben. A továbbiakban egy-egy AVS bemutatásakor csak az újdonságokra fogunk kitérni.

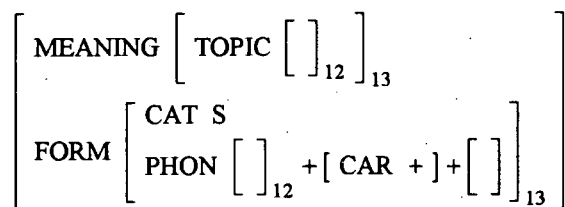
(12) *Igei csoport*



Egy igei csoport jelentését tehát úgy kapjuk meg, hogy az ige jelentését mint függvényt alkalmazzuk a tárgy jelentésére. Szintaktikailag az igei csoport egy igéből és egy főnévi csoportból áll, amelyek helyzete egymáshoz képest nem meghatározott, és a mondat többi elemével is keveredhetnek (egyszerű kapcsos zárójel). Végezetül az igei csoport INFL jegyének értékét az ige INFL jegye határozza meg, így ezt a jegyet örökli az egész mondat az előző szabály értelmében.

A semleges mondatok felépítésében a legfontosabb szabály a következő, hiszen mint mondtuk, a legfontosabb az aktuális mondattagolás ezen mondatok felszíni szerkezetének meghatározásában, márpedig a következő szabály éppen azt mondja ki, hogy a mondat topikja mindig közvetlenül azon elem előtt áll, amelyiknek CAR jegye a „+” értéket veszi fel, vagyis – mint azt már fentebb említettük – amelyik a komment elején áll.⁶ Ezek megelőzik a mondat azon maradvékát, amelyet nem kényszerít valamilyen más szabály a mondat élén állni, amire láttunk már példát (ld. (1g)).

(13) *Topik-komment szerkezet*



⁶Mint említettük, sok mindent leegyszerűsíthetünk. Megjegyzendő például, hogy a komment elején még ún. kvantoros kifejezések (pl. *mindig*, *mindenkinek* stb.) is állhatnak.

Ez a három szabály már elég is lesz, hogy az (1a) példát illusztrációképpen levezessük. Mint fent említettük, kiindulásul egy olyan jelentés-reprezentáció szolgál, amely a beszélő szándékát fejezi ki, hogy tudniillik a már említett Mariról meg kívánja jegyezni, hogy látta Jánost. Ezt fejezi ki az alábbi parciális (mert csak a jelentésre vonatkozó információt tartalmazó) jegy-érték struktúra:

(14) 'Mari látta Jánost'

$$\left[\begin{array}{l} \text{MEANING} \\ \left[\begin{array}{l} \text{SUBJ } \boxed{1} \text{ Mary}_1 \\ \text{PRED } \left[\begin{array}{l} \text{FUN saw}_2 \\ \text{ARG John}_3 \end{array} \right]_4 \\ \text{TOPIC } \boxed{1} \end{array} \right]_5 \end{array} \right]$$

Ez az AVS unifikálható a fenti három AVS közül kettővel, vagyis a beszélő ezeket a szabályokat alkalmazhatja, hogy kifejezze a szándékolt jelentést:

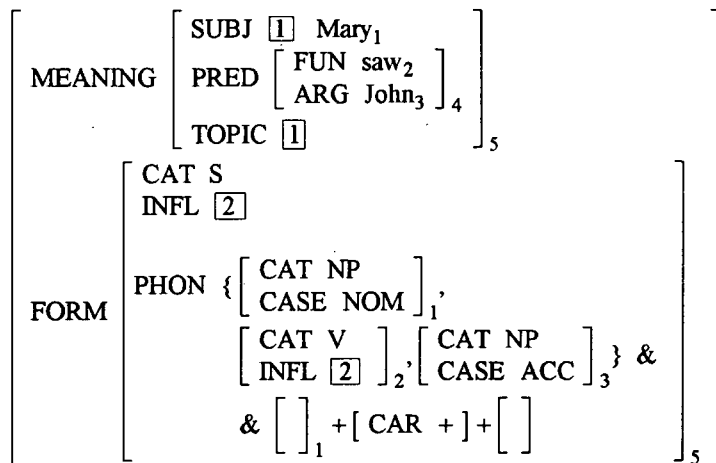
(15) = (14) \sqcup (11) \sqcup (13)

$$\left[\begin{array}{l} \text{MEANING} \\ \left[\begin{array}{l} \text{SUBJ } \boxed{1} \text{ Mary}_1 \\ \text{PRED } \left[\begin{array}{l} \text{FUN saw}_2 \\ \text{ARG John}_3 \end{array} \right]_4 \\ \text{TOPIC } \boxed{1} \end{array} \right]_5 \\ \\ \text{FORM} \\ \left[\begin{array}{l} \text{CAT S} \\ \text{INFL } \boxed{2} \\ \\ \text{PHON } \left\{ \left[\begin{array}{l} \text{CAT NP} \\ \text{CASE NOM} \end{array} \right]_1, \left[\begin{array}{l} \text{CAT VP} \\ \text{INFL } \boxed{2} \end{array} \right]_4 \right\} \& \\ \& \left[\right]_1 + \left[\text{CAR } + \right] + \left[\right] \end{array} \right]_5 \end{array} \right]$$

A PHON jegy értéke (illetve az „&” jel) azt mondja, hogy a felszíni szerkezetnek egyszerre két megkötésnek kell megfelelnie, amelyeket a két egyidejűleg megvalósítható szabály megkövetel, ahogy azt fent részletesen elmagyaráztuk. Továbbá felhívjuk a figyelmet, hogy a szemantikai koindexálás miatt az unifikációnál az alany és a topik szintaktikai megfelelőjének indexe egybeesett, vagyis itt derül ki, hogy ennek a két szemantikai elemnek a szintaktikai megfelelője is egy és ugyanaz az elem.

A következő lépésben az igei csoportot bontjuk ki, megint a szemantikai reprezentáció által vezérelve:

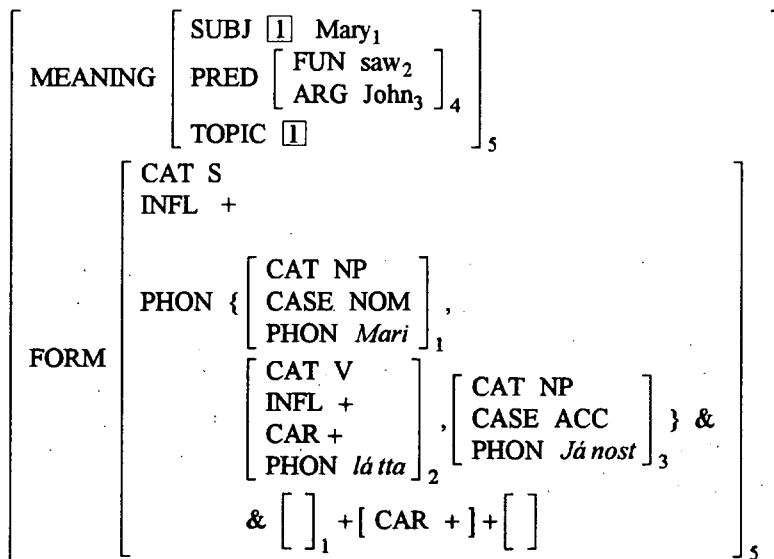
(16) = (12) beépítése (15)-be a 4. indexnél



Most láthatunk példát arra, hogy a kapcsos zárójelben szereplő kapcsos zárójel eltűnik, vagyis az igei csoport elemei immár nemcsak egymással, de a mondat többi elemével is szabadon keveredhetnek. Ez a reprezentáció mutatja azt is, hogy hogyan örökli a mondat az ige INFL jegyének értékét.

Már csak egy lépés maradt hátra, tudniillik hogy a lexikai reprezentációkat beépítsük a fenti konstrukcióba, amit most egy lépésben megtehetünk:

(17) = (7) beépítése (16)-ba az 1. indexnél, (6) beépítése (16)-ba a 3. indexnél, (8) beépítése (16)-ba a 2. indexnél



Láthatjuk, hogy az *látta* igének megfelelő reprezentáció beépítésével kapja meg a mondat INFL jegye a „+” értéket. Az eredményre már nem alkalmazható egyetlen szabály sem, de erre nincs is szükség, mert a PHON jegy megfelelő megszorításokat tesz a felszíni szerkezetre, tudniillik hogy szerepeljen benne a *Mari*, a *Jánost* illetve a *látta* jelsorozat, mégpedig úgy, hogy az alany álljon a mondat elején, közvetlenül az ige előtt, márpedig ezt kizárólag az (1a) mondat elégíti ki. (Itt és a továbbiakban feltételezzük, hogy a többi szó [CAR-] jegyű, ellenkező esetben a sztenderd unifikációs mechanizmus más sorrendet is megengedne.)

Világos, hogy ugyanezek a szabályok megfelelő kiindulójelentés-reprezentáció mellett elszámolnak az (1b) példával is. A (c) példa leírásához sem lesz további eszközre szükségünk, csak a kiindulójelentés-reprezentációt kell megfelelően megválasztanunk. Ez a következő lesz:

(18) ‘Látta Mari Jánost.’

$$\left[\text{MEANING} \left[\begin{array}{l} \text{SUBJ } \text{Mary}_1 \\ \text{PRED} \left[\begin{array}{l} \text{FUN } \text{saw}_2 \\ \text{ARG } \text{John}_3 \end{array} \right]_4 \\ \text{TOPIC } - \end{array} \right] \right]_5 \right]$$

Ez a szemantikai reprezentáció valóban azt kódolja, amit mondani szeretnénk, vagyis azt a nagyon jelölt esetet, amikor a szándékolt közlés beszédértéke semleges, de nem kapcsolódik semmilyen előzményhez sem, azaz üres topikot tartalmaz.⁷ Kiszámítható, hogy a megfelelő szabályok ((11), (12) és (13)) alkalmazása az előbbi példa kapcsán bemutatott módon valóban a kívánt felszíni sorrendet hozza létre, pontosabban egy olyan reprezentációt, amely csak azt köti meg, hogy az igének (vagy ha valami más viseli a CAR jegyet, mint az *elütötte* szóban az igeikötő, akkor annak) kell a mondat élén állnia, míg a többi elem sorrendje szabad, tehát mivel az (1c) példában szereplő mindkét mondat kielégíti ezt a feltételt, grammatikánk mindkettőt generálja, azonos jelentéssel.

Mielőtt rátérnénk az irtó prozódiajú mondatok elemzésére, ki kívánunk térni az igeikötők elszakadásának problémájára, és ezen belül is a segédige beférkőzésére az ige és igeikötője közé, amit a (19) példa illusztrál. Itt kap különleges szerepet az igevivő-ségre utaló CAR jegy, ezért most részletesebben is szólnunk róla.

A legtöbb igének *igevivőre* van szüksége ahhoz, hogy semleges (nem *folyamatos*, vagyis progresszív) igeszemléletű kommentben szerepeljen. Az azonban, hogy milyen szó vagy szerkezet az igevivő (illetve hogy szükség van-e rá egyáltalán), az ige szótári sajátossága, ezért ezt a szabályszerűséget nem írhatjuk fel általános konstruk-

⁷Itt és a továbbiakban is feltételezzük, hogy a „-” érték a [] AVS-en kívül semmilyen más értékkel sem unifikálható.

cióként. Bizonyos igék a helyhatározót kívánják igeivőnek (*a Tisza két partján terül el*), mások az alanyt (*homok ment/jutott/került a szemembe*), megint mások a célhatározót (*Szegedre utazik*), egyesek a tárgyat (*fát vág*), némelyik ige bármilyen infinitívuszt (*hinni/eltűnni látszik*), végül némelyik egyáltalán nem kíván igeivőt (*utál futballozni*). Az igeiktős igéknek általában az igeiktő az igeivőjük (*me nézi a filmet*), ha pedig valamelyiknek nincs szüksége igeivőre, akkor azt mondhatjuk, hogy ő maga az igeivő (*úszik*).

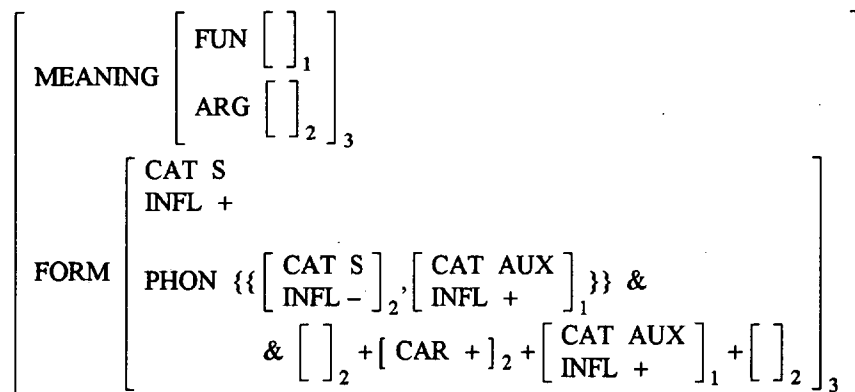
Az igeivőség legmegbízhatóbb tesztje, hogy az eldöntendő kérdésre helyeslőleg (*Igen helyett*) „rövid válaszként” mindig az igeivőt lehet válaszolni. Az igeivők kezelésének leghelyesebb módja az, ha a szótárban minden egyes igeosztálynál egy általános konstrukcióval megadjuk, hogy milyen szónak vagy morfémanak ítélik oda az illető osztály igéi, az igeivői szerepet, mit látnak el a [CAR +] jegy-érték párral. Ezt a tulajdonságot azután az osztály (nem rendhagyó) tagjai öröklék ettől az általános konstrukciótól a konstrukciórendszer hierarchikus felépítése miatt, amire itt nem térünk ki.

Csak egy alosztályra fogunk nagyobb figyelmet fordítani az alábbiakban, a *segédigék*re. Ezekre az a jellemző, hogy *közvetlenül maguk elé követelik a nekik alárendelt ige igeivőjét* (innen a „beférkőzés” is):

a Tisza két partján fog elterülni
homok fog a szemembe kerülni
Szegedre akar utazni
fát szeretne vágni
meg tudja nézni a filmet
úszni akar

A semleges mondatnak tehát egy sajátos esete a segédigés, amelyben a segédigét közvetlenül meg kell előznie az argumentuma igeivőjének:

(20) *Segédigés szerkezet*



Azt, hogy ez egy mondat szintű konstrukció, a segédigék szemantikája bizonyítja, hiszen ezek mondatoperátorok, ahogy ezt a fenti reprezentáció is mutatja.⁸ A FORM jegy értéke a következő megszorításokat teszi: az egész konstrukció kategóriája finit mondat, amely áll egy infinitívuszi mondatból és egy ragozott segédigéből, mégpedig úgy, hogy a beágyazott mondat igevivője fogja a segédige előtti pozíciót kitölteni, míg a beágyazott mondat többi része a más alkalmazandó szabályok megkövetelte módon előzi meg illetve követi őket.⁹

Már csak a *fogja* szó szótári alakját kell megadnunk, hogy levezethessük a (3) példát:

(21) *fogja*

$$\left[\begin{array}{l} \text{MEANING } \text{will}_{16} \\ \text{FORM } \left[\begin{array}{l} \text{CAT } \text{AUX} \\ \text{INFL } + \\ \text{PHON } \textit{fogja} \end{array} \right]_{16} \end{array} \right]$$

Most már minden eszköz a rendelkezésünkre áll, hogy levezessük a kívánt mondatot, amit illusztrációképpen az alábbiakban meg is teszünk. Levezetésünk a következő szemantikai reprezentációból indul ki:

(22) 'Jánost el fogja ütni a vonat'

$$\left[\begin{array}{l} \text{MEANING } \left[\begin{array}{l} \text{FUN } \text{will}_1 \\ \text{ARG } \left[\begin{array}{l} \text{SUBJ } \textit{the train}_2 \\ \text{PRED } \left[\begin{array}{l} \text{FUN } \textit{hit}_3 \\ \text{ARG } \boxed{1} \textit{John}_4 \end{array} \right]_5 \end{array} \right]_6 \\ \text{TOPIC } \boxed{1} \end{array} \right]_7 \end{array} \right]$$

Az első lépésben ez a reprezentáció unifikálható a segédigét tartalmazó mondatok általános sémájával:

⁸A kettős kapcsos zárójel itt az a triviális követelményt hivatott biztosítani, hogy különböző mondatok elemei nem keveredhetnek. Míg pl. a (11) konstrukcióban egyszeres kapcsos zárójelet használtunk, hiszen alany-állítmány szerkezetek nemcsak önálló mondatként fordulhatnak elő, itt biztosítani akartuk, hogy a finit segédigés mondatok önálló megnyilatkozást alkossanak.

⁹A FORM jegy értékében két követelmény itt azt fejezi ki, amit a GPSG-ben és a hasonló elméletekben ún. közvetlen dominancia (ID) ill. lineáris sorrend (LP) szabályokkal lehetne felírni.

(23) = (22) \sqcup (20)

MEANING	$\left[\begin{array}{l} \text{FUN will}_1 \\ \text{ARG} \left[\begin{array}{l} \text{SUBJ the train}_2 \\ \text{PRED} \left[\begin{array}{l} \text{FUN hit}_3 \\ \text{ARG} \boxed{1} \text{ John}_4 \end{array} \right]_5 \end{array} \right]_6 \end{array} \right]_7$
FORM	$\left[\begin{array}{l} \text{CAT S} \\ \text{INFL +} \\ \text{PHON} \{ \{ \left[\begin{array}{l} \text{CAT S} \\ \text{INFL -} \end{array} \right]_6, \left[\begin{array}{l} \text{CAT AUX} \\ \text{INFL +} \end{array} \right]_1 \} \} \&$
	$\& \left[\right]_6 + \left[\text{CAR +} \right]_6 + \left[\begin{array}{l} \text{CAT AUX} \\ \text{INFL +} \end{array} \right]_1 + \left[\right]_6 \left. \right]_7$

Ezután két lépést összevonva ezt a reprezentációt a topik-komment szerkezet kifejezésének céljából unifikáljuk a (13) konstrukcióval, ill. az alany-állítmány szerkezet kifejezéséhez beleépítjük a (11) konstrukciót a megfelelő indexnél:

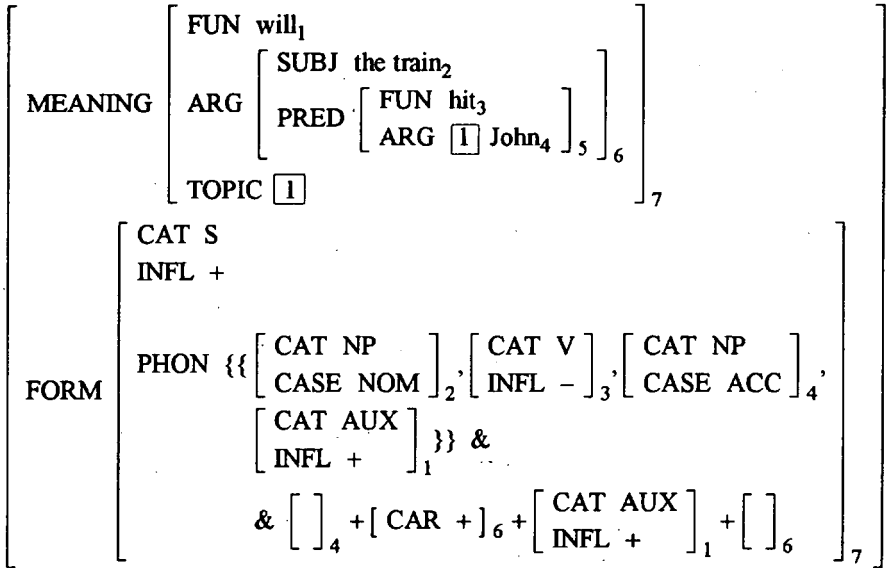
(24) = (11) beépítése (23) \sqcup (13)-ba a 6. indexnél

MEANING	$\left[\begin{array}{l} \text{FUN will}_1 \\ \text{ARG} \left[\begin{array}{l} \text{SUBJ the train}_2 \\ \text{PRED} \left[\begin{array}{l} \text{FUN hit}_3 \\ \text{ARG} \boxed{1} \text{ John}_4 \end{array} \right]_5 \end{array} \right]_6 \end{array} \right]_7$
FORM	$\left[\begin{array}{l} \text{CAT S} \\ \text{INFL +} \\ \text{PHON} \{ \{ \left[\begin{array}{l} \text{CAT NP} \\ \text{CASE NOM} \end{array} \right]_2, \left[\begin{array}{l} \text{CAT VP} \\ \text{INFL -} \end{array} \right]_5, \left[\begin{array}{l} \text{CAT AUX} \\ \text{INFL +} \end{array} \right]_1 \} \} \&$
	$\& \left[\right]_4 + \left[\text{CAR +} \right]_6 + \left[\begin{array}{l} \text{CAT AUX} \\ \text{INFL +} \end{array} \right]_1 + \left[\right]_6 \left. \right]_7$

Mint látható, a (11) konstrukciónak megfelelően a PHON jegy értékében szereplő nem finit mondat helyére az alany (indexe 2) és az állítmány (indexe 5) került, és az igei csoport örökölte az INFL jegy értékét. A (13) konstrukcióval való unifikációnak annyi a hatása, hogy a [CAR +] előtti hely indexét egybeejtjük a topik indexével, ami 4, hiszen a topik értéke azonos a MEANING ARG PRED ARG értékével.

A következő lépésben kifejtjük a VP tartalmát a (12) konstrukció alkalmazásával:

(25) = (12) beépítése (24)-be az 5 indexnél



Ettől a pillanattól fogva a leírás formai részéből is kiderül, hogy a segédigének alárendelt tárgy a mondat topikja.

Ezek után már csak a lexikai elemek behelyettesítése van hátra:

(27) *Fókuszos mondat*

$$\left[\begin{array}{l} \text{MEANING} \left[\begin{array}{l} \text{FOCUS []}_i \\ \text{BACKGROUND []}_j \end{array} \right]_k \\ \text{FORM} \left[\begin{array}{l} \text{CAT S} \\ \text{INFL +} \\ \text{PHON []}_i + [\text{CAT V}] + [\]_j \end{array} \right]_k \end{array} \right]$$

(28) *Nyomatékos mondat*

$$\left[\begin{array}{l} \text{MEANING [EMPHATIC []]}_i \\ \text{FORM} \left[\begin{array}{l} \text{CAT S} \\ \text{INFL +} \\ \text{PHON [CAR +]}_i + [\]_i \end{array} \right]_k \end{array} \right]$$

Mivel mindkét konstrukció mondat szintű (az egész mondat szemantikája sajátos), felülbírálják az alacsonyabb rendű szórendi előírásokat. A fókuszos mondat szórendi sajátosságai miatt az igeívő nem állhat az ige előtt, rendes helyén (hacsak nem ő maga a fókusz); a nyomatékos mondatnál rendszeren az igeívő kapja az irtó hangsúlyt (tehát beférkőzés is van), ha nem, akkor nem szokványos nyomatékos mondatról van szó (például az ún. egzisztenciális mondatban a ragozott igeívő kapja a hangsúlyt – ez a progresszív mondatok irtó hangsúlyú megfelelője).

A másik jelenség, amelyet az irtó prozódiajú mondatoknál megfigyelhetünk, a *kontrasztív topik* esetleges jelenléte, a topikhoz hasonló helyzetben. Ennek sajátos jelentéstartalma van (azt fejezi ki, hogy a kontrasztív topikra vonatkozó információ, ami a mondatban van, nem kimerítő), tehát nem azonos a semleges mondat topikjával, kiválasztásának elvei sem azonosak. Ezért ennek az elhelyezkedése is felülbírálja az alacsonyabb szintű szórendi követelményeket. Sematikus ábrázolása:¹⁰

(29) *Kontrasztív topik*

$$\left[\begin{array}{l} \text{MEANING} \left[\begin{array}{l} \text{CONTR_TOPIC []}_i \\ \text{REMAINDER []}_j \end{array} \right]_k \\ \text{FORM} \left[\begin{array}{l} \text{CAT S} \\ \text{PHON []}_i + [\]_j \end{array} \right]_k \end{array} \right]$$

¹⁰Eltekintünk itt azoknak a mondatoknak a problémájától, amelyekben a kontrasztív topik a fókuszos összetevő vagy az ige másolata: 'Enni János evett; 'Ételnek 'étel volt.

Ez a konstrukció persze kissé triviális. A részletes jelentéstani jellemzés persze igen bonyolult lenne, de a formai követelmények a valóságban is triviálisak. Úgy is mondhatnánk, hogy a konstrukciós nyelvtan módszerével frappáns leírását adhattuk annak a ténynek, hogy a kontrasztív topik mindig a mondat élén áll, akármilyen is a mondat logikai szerkezetében betöltött szerepe.

IRODALOM

- Fillmore, Charles J. – Paul Kay 1993: *Construction grammar coursebook*, Berkeley, University of California.
- Fillmore, Charles J. – Paul Kay – Mary Catherine O'Connor 1988: Regularity and idiomaticity in grammatical constructions: The case of let alone, *Language* 64, 501–538.
- Hockett, Charles F. 1958: *A course in modern linguistics*, New York, Macmillan.
- Johnson, Mark 1988: Attribute-value logic and the theory of grammar, CSLI Lecture Notes.
- Kálmán C. György – Kálmán László – Nádasdy Ádám – Prószezy Gábor 1989: A magyar segédigék rendszere, *Általános nyelvészeti tanulmányok XVII*, 49–103.
- Manaster-Ramer, Alexis – Wlodek Zadrozny 1992: Open idioms and constructions, *Proc. Conf. on Idioms*, Tilburg, ITK.
- Zwicky, Arnold 1989: Idioms and constructions, *ESCOL* 5, 547–558.

DISCONTINUOUS CONSTITUENTS IN HUNGARIAN FROM A CONSTRUCTIONIST POINT OF VIEW

LÁSZLÓ KÁLMÁN – GÁBOR RÁDAI

Budapest University, Theoretical Linguistics Department

c/o Research Institute for Linguistics

of the Hungarian Academy of Sciences

Budapest I., Színház u. 5–9. P.O. Box 19. H-1250

E-mail: kalman@nytud.hu, radai@nytud.hu

Discontinuous constituents are notoriously difficult to treat for any generative grammar and especially so in the case of Hungarian, where the word order of sentences depends on the communicative role of the constituents rather than their grammatical role. We argue that construction grammar is highly adequate for the description of these phenomena, since it is exactly the fact that different constructions are realised simultaneously in a Hungarian sentence – such as, e.g., the subject verb, the topic comment, the focus background or the contrastive topic construction – that is responsible for word order. We represent constructions and lexical entries by Attribute Value Structures, which encode both syntactical and semantic information, and the generation of a sentence is driven by the semantic representation, in such a way that all syntactic constraints are satisfied by the sentence. A further claim is that an additional feature needs to be encoded, namely constructions often assign a particular position to an unspecified constituent marked ‘carrier +’ the identity of which depends on the verb., i.e. it is specified by the AVS corresponding to it.