

Megjegyzések a grammatikalitás fokairól

A grammatikaelmélet fontos feladatai közé tartozik a grammatikalitás fokainak megállapításához szükséges módszerek kidolgozása. Közismert tény, hogy a megértett 'mondatok' száma lényegesen nagyobb, mint a grammatikailag (és szemantikailag) helyes mondatok száma. Erre a jelenségre magyarázatot várunk a grammatikaelmélettől. Pontosabban a következő kérdések megválaszolását is megkívánjuk a grammatikaelmélettől:

a) Milyen mértékben térhet el a grammatikailag helytelen mondat grammatikailag helyes párjától?

b) Hogyan lehet az eltérés fokát megállapítani?

c) Hol az a határ, ahol a „mondat” már érthetelenné válik?

Ezek a kérdések nem újak, hiszen — mint ismeretes — CHOMSKY a grammatikalitás fokaival külön cikkben foglalkozik.¹ Ami azonban az egész generatív grammatika felépítésével kapcsolatban igaz, az különösen szembevető a grammatikalitás fokaival kapcsolatban: tudniillik az a tény, hogy CHOMSKY és tanítványai elsősorban az angol és az angolhoz közelálló nyelvek vizsgálatából szűrték le általános tapasztalataikat, de egyetlen agglutináló nyelvet sem vizsgáltak meg részletesen. Ebből a tényből következik, hogy a generatív grammatikát több szempontból módosítani és kiegészíteni kell ahhoz, hogy agglutináló nyelvek esetében is a várt eredményt adja. Ezzel a kérdéssel azonban nem kívánunk most részletesen foglalkozni.² Bennünket most CHOMSKY-nak a grammatikalitás fokaival kapcsolatos elgondolásainak kiegészítése érdekel. Ahhoz, hogy CHOMSKY elméletének kiterjesztésével foglalkozhassunk, röviden össze kell foglalnunk ezen elmélet legfontosabb elgondolásait.

CHOMSKY abból a feltételezésből indul ki, hogy adott a kategóriáknak egy hierarchikus rendszere, úgyhogy az első szinten egyetlen kategóriánk van: C_1^1 (amely például minden szónak lehet kategóriája); a második szinten a következő kategóriáink vannak: $C_1^2, \dots, C_{n_2}^2$; a harmadik szinten: $C_1^3, \dots, C_{n_3}^3$, ahol $n_3 > n_2$ és így tovább, amíg el nem érünk az m -edik szintig, ahol a kategóriáink: C_1^m, \dots, C_n^m ($1 < n_2 < n_3 < \dots < n_m$). Mindegyik szinten a kategóriák az egész szókészletet felölelik, azaz mindegyik szó legalább egy kategóriához tartozik (homonimia esetén többhöz is). A kategóriáknak ilyen rendszere

¹ CHOMSKY, N., Degrees of Grammaticalness, The Structure of Language, Readings in the Philosophy of Language (J. A. FODOR, J. J. KATZ, eds.), Prentice-Hall, Inc. Englewood Cliffs, New Jersey (1964), 384–389. o.

² Ezzel a problémával foglalkozott — többek között — az MTA Nyelvtudományi Intézetében 1965 júniusában megtartott „A magyar nyelv generatív nyelvtana” című munkaértekezlet is.

tulajdonképpen azt jelenti, hogy mindegyik szint finomabb felosztása az előbbinek, a kategóriák alkategóriákra bomlanak, majd utóbbiak ismét alkategóriákra és így tovább.³

Példaképpen tekintsünk egy három szintű hierarchiát. C_1^1 legyen az összes szavak osztálya, továbbá C_1^2 a főnevek, C_2^2 az igék, C_3^2 a melléknevek és C_4^2 a többi szó osztálya. Jelöljük C_1^3, \dots, C_j^3 -mal az igék alosztályait (intranszítív igék, tranzitív igék élő tárgygal stb.), a főnevek alosztályait (élő, élettelen stb.) és egyéb alosztályokat. Minden szó-lánc a fentiek alapján első szintű, második szintű és harmadik szintű kategória láncsal jellemezhető. Például a

(1) Péter szereti a magányt

a következőképpen jellemezhető:

Első szinten: $C_1^1 C_1^1 C_1^1 C_1^1$

Második szinten: $C_1^2 C_2^2 C_4^2 C_2^2$ (azaz NVAN)

Harmadik szinten: $N_{\text{élő}} V_{\text{tr, x}} A N_{\text{absztrakt}}$

ahol x azt jelenti, hogy a szóban forgó V_{tr} (= szeret) csak élő alannyal állhat. Hasonlítsuk mármost össze az (1) mondatot a következő szó-láncokkal:

(2) Asztal szereti a magányt.

(3) Piros szereti a magányt.

A (2) szó-lánc az első két szinten ugyanazzal a kategória-lánccal jellemezhető, mint (1). Eltérés csak a harmadik szinten van: „asztal” nem élő alany. (2) esetében viszont már a második szinten sincs egyezés (az átvitt értelmezéstől eltekintünk), hiszen „piros” kategóriája nem N. Ez más szóval azt jelenti, hogy (3) jobban eltér a grammatikailag helyestől, azaz (1)-től, mint (2). (2) csak a harmadik szinten tér el, tehát közelebb áll (1)-hez, mint (3), amely már a második szinten eltér. Ily módon a grammatikailag helytelen mondatok osztályozásához juthatunk és minél finomabb a kategória-rendszerünk (azaz minél több szintből áll a hierarchiánk), annál finomabb különbséget tudunk tenni az eltérések között. A grammatikailag helyestől való eltérés annál nagyobb, minél magasabb szintű kategóriában van eltérés.

CHOMSKY máshelyütt azt a fontos megjegyzést teszi,⁴ hogy azok a grammatikailag helytelen mondatok, amelyek csak a szelekciós szabályok megsértése miatt helytelenek (mint a mi fenti példánkban (2)) általában könnyen értelmezhetők metaforikusan, míg azok, amelyek szófaji kategóriákban térnek el, általában teljesen értelmetlenek. Véleményünk szerint a magyarban nem egészen ez a helyzet, mint ahogyan az a fenti példából is kiviláglik: (3) könnyen értelmezhető. Ahhoz azonban, hogy ezt a problémát világosan lássuk,

³ Az ilyen kategória-rendszer részletesebb jellemzésével másutt foglalkozunk: Similarity, Contrast and Inclusion in Natural Language, Linguistics (megjelenés alatt).

⁴ CHOMSKY, N., Categories and Relations in Syntactic Theory, Kézirat (MIT, Cambridge, Mass.), 1964. Hasonló megjegyzést találunk már BLOOMFIELDNÁL is: Language, New York, Holt, 1933.

szükségünk lenne a megfelelő hierarchikus kategória-rendszerre. Erről a rendszerről azonban in concreto alig tudunk még valamit.

CHOMSKY módszere kétségtelenül helyes, de csak részben ad feleletet a feltett kérdésekre. Most már látjuk, hogy az eltérés fokát milyen módszerrel lehetne megállapítani, és azt is, hogy milyen módon lehetne megpróbálkozni az érthetetlen és az érthető közti határ megvonásával. Mindezeket a kérdéseket a fent vázolt módszerrel meg lehet válaszolni. Hátra van még az a) kérdés megválaszolása.

a) esetén arról van szó, hogy egy-egy grammatikailag helytelen 'mondat'⁵ csak grammatikailag helyes párja alapján érthető meg. Azaz például

Pali játszik a kertbe

megértése csak a

Pali játszik a kertben

alapján történhet. Ezt nevezhetjük analógia alapján történő megértésnek. Ezzel a kérdéssel Chomsky nem foglalkozik.

A b) és a c) kérdések esetén viszont az értelmezés, amennyiben egyáltalán lehetséges, nem analógia alapján történik, hiszen a jelentés lényegesen megváltozhat, szavak kicserélése esetén az eredeti mondat egyáltalán nem rekonstruálható. (2) eredeti mondata mi lenne? Annyit tudunk, hogy „asztal” helyett élő alanyt kell vennünk. Más szóval, mondatoknak egy osztályát kapjuk. És ez érvényes általánosan is. A CHOMSKY-féle grammatikalitási fokok esetén az eltérés sohasem egy adott mondatról értendő, hanem mondatoknak egy osztályától. Éspedig minél nagyobb az eltérés (a fenti értelemben), annál tágabb ez az osztály.

Bár a CHOMSKY-féle kategória-hierarchia kiterjeszhető toldalékmorfémák esetére is, azaz olyan jellegű eltérések is megmagyarázhatók, mint

- (4) Péter esz káposzta.
Virág nyílni az udvarba.
Asztalos asztal csinálni.

De ilyen esetekben más megfontolások is érvényesek, amelyek megmutatják, hogy CHOMSKY elgondolásai milyen mértékben és milyen irányban szorulnak kiegészítésre.

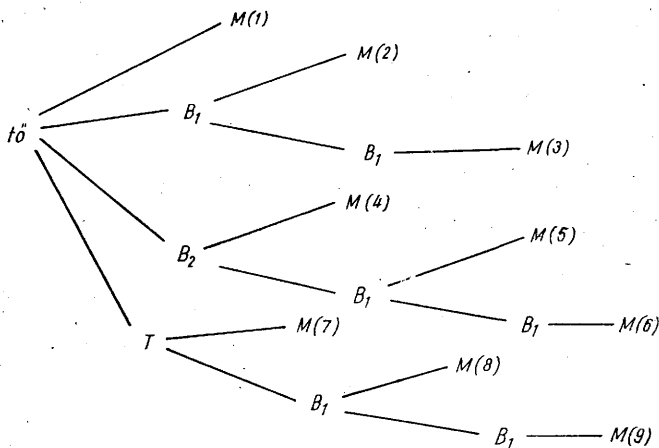
Először is nyilvánvaló, hogy (4) mondataink esetén a megértés analógia alapján történik és könnyű rekonstruálni a megértés alapját képező mondatot (általában egyetlen mondatot).

Mielőtt azonban a felvetett kérdésre válaszolnánk, meg kell vizsgálnunk a magyar szóalakok generálásának kérdését. Az agglutináló nyelvek esetén a szó-generálás lényegesen fontosabb és egyben bonyolultabb probléma, mint az angolhoz hasonló nyelvek esetén. A szóalakok generálására azonban a transzformációs modellnél lényegesen gyengébb modell is elegendő, mégpedig az

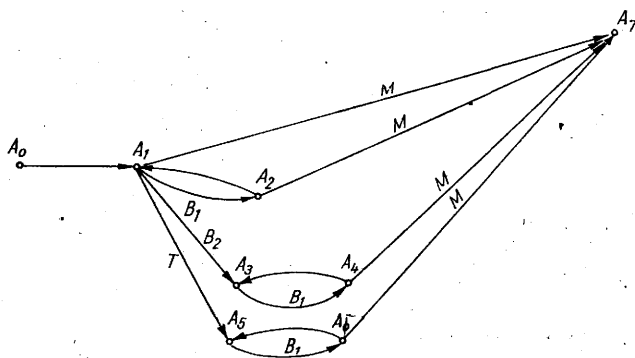
⁵ 'Mondat' helyett itt tulajdonképpen 'szó-láncot' kellene használnunk. A továbbiakban — a pontosabb megfogalmazás kedvéért — szó-lánccról beszélünk akkor, amikor vagy nem tudjuk, hogy helyes-e grammatikailag a 'mondat', vagy pedig tudjuk, hogy nem az.

úgynevezett véges állapotú grammatika.⁶ Az a. ábra a generálás diagramját, a b. ábra a véges állapotú grammatikát mutatja.

Jelmagyarázat: B_1 a következő morfémákat tartalmazza: é, éi; T a többes szám végződését, B_2 a birtokos ragokat, M pedig a végződéseknek egy tágabb osztályát (tárgy, határozói, és a \emptyset végződést stb.) tartalmazza.



a.



b.

⁶ A véges állapotú grammatikát részletesen ismertetem Matematikai nyelvészeti tanulmányok (OMKDK, 1964) című könyvemben.

⁷ Néhány megjegyzést kell tennünk az ábrához. Az M mellett zárójelben álló számok az ágak megkülönböztetésére szolgálnak. Legyen most a 'szék' az adott tő. Az (1) ág a következő alakok generálását mutatja: 'szék', 'széket', 'széknek', 'székkel' stb. A (2) ág: 'széké', 'székét', 'székének', 'székével', ..., 'székéi', 'székéit', 'székéinek', 'székéivel', ... A (3) ág: 'székéé', 'székéét', 'székéének', 'székéével', ..., 'székééi', 'székééit', 'székééinek', 'székééivel', ..., 'székééie', 'székééiét', 'székééiének', 'székééiével', ..., 'székééiei', 'székééiét', 'székééiének', 'székééiével' ... A (4) ág: 'székem', 'székemet', 'székemnek', 'székemmel', ..., 'székeim', 'székeimet', 'székeimnek', 'székeimmel', ... Az (5) ág: 'székemé', 'székemét', 'székemének', 'székemével', ..., 'székeimé', 'székeimét', 'székeimének', 'székeimével', ... A (6) ág: 'székeméé', 'székeméét', 'székeméének', 'székeméével', ...

Hasonló modell dolgozható ki a többi szófaj számára is. Mármost agglutináló nyelvek esetén egy-egy morféma megváltozása vagy elhagyása még nem jelenti a grammatikalitás megsértését. Általános érvényű a következő kritérium: Az eredeti mondat csak abban az esetben rekonstruálható, ha valamely morféma megváltozása vagy elhagyása esetén grammatikailag helytelen mondathoz jutunk. Ez azonban csak szükséges és nem elegendő feltétel. Világítsuk meg ezt néhány példával.

Nyilvánvaló, hogy olyan szó-láncok esetén, mint (4) a rekonstrukció megvalósítható, de nem valósítható meg a következő esetekben:

Szekrényeimben nincs hely.

Szekrényemben nincs hely.

Szekrényben nincs hely.

Szekrényeiben nincs hely.

Szekrényében nincs hely.

...

stb.

Ilyen esetekben legfeljebb a beszéd-szituációból következtethetnénk arra, hogy valami nincs rendjén, ami viszont extralingvisztikai vonatkozás és mint ilyen nem tartozhat vizsgálódásaink körébe.

Másrészt a feltétel nem elegendő, mert például

Péter hoz autót.

Az iskolában nem kapta könyv.

Öreg ember nem jó veszekedni.

Ezek természetesen egyértelműek lehetnek a beszéd-szituáció ismeretében, formálisan tekintve azonban mindegyiküknek mondatoknak egy osztálya felel meg, amelyek a CHOMSKY-féle kategória-hierarchia esetétől eltérően mind különböző grammatikai struktúrájú mondatból állnak. Azaz például első szó-láncuk esetén arra is gondolhatnánk, hogy „Péter hoz autót”, de arra is, hogy „Pétert autó hozza” stb.

Visszatérve a főnév-alakokat generáló modellünkhöz, általában megállapítható, hogy a B_1 elhagyása vagy a B_2 elemeinek egymás közti felcserélése, illetőleg elhagyása nem rekonstruálható. (Természetesen hangsúlyos eset-

nek', 'székeméivel', ..., 'székeiméé', 'székeiméét', 'székeiméének', 'székeiméével', ...
A (7) ág: 'székek', 'székeket', 'székeknek', 'székekkel', ... A (8) ág: 'székeké', 'székekét', 'székekének', 'székekével', ..., 'székekéi', 'székekéit', 'székekéinek', 'székekéivel', ...
A (9) ág: 'székekéé', 'székekéét', 'székekéének', 'székekéével', ..., 'székekéié', 'székekéiét', 'székekéiének', 'székekéiével', ...

Megjegyezzük, hogy a B_1 osztály képviseli a rekurzív elemet a generálásban. Az ábrán csak kétszer fordul elő egy ágban, elvileg azonban akárhányszor előfordulhat. Ez kivüláglik a b) ábrából, ahol ez a tulajdonság 'hurok' formájában jut kifejezésre. A b) ábrán megadott véges állapotú grammatika szerint helyes a következő is:

asztaléiéi ... székekéiéi ... nek

Határt tulajdonképpen csak pszichológiai megfontolások alapján szabhatunk. Grammatikailag csak ezért tartjuk helytelennek a fenti alakokat, mert érthetetlenek, azaz kimondottan pszichológiai okokból.

ben ez nem áll. Pl. Az én házban ilyen nem szabad csinálni.) A T osztály esetén mind az elhagyás, mind az egymás közti felcserélés megsérti a grammatikalitást, tehát a rekonstruálhatóság szükséges feltétele teljesül. Ha csak felcserélésről van szó, pl. „székok” „székek” helyett, akkor minden esetben rekonstruálható, elhagyás esetén pedig csak akkor, ha az ige alakja lehetővé teszi. Például:

Az ember sokat dolgoznak.

De ha csak ennyi állna:

Az ember sokat dolgoz.

a rekonstrukció már nem lenne egyértelmű. Ez is azt mutatja, hogy az igének központi jelentősége van a rekonstrukciónál. Tárgy helyzetben azonban T ismét nem rekonstruálható általában.

Például:

Az író regény ír.
Az emberek ház építenek.

(regényt vagy regényeket? házat vagy házakat?).

Ami az M osztályt illeti, ez az osztály nem egységes. A tárgyrag rekonstruálható az ige helyes alakjának ismeretében, ugyanez vonatkozhat a részes-határozói rag esetére, bár a kettős tárgy már problémát okozhat. Például:

(5) Adok alma Péter.
Építünk magunk ház.

és

(6) Szerzünk csónak ház.
Keresünk egér macska.

(5) esetén a grammatikailag helyes mondat rekonstruálható, (6) esetén nem. Általában igaz, hogy az igealak helyessége mellett a szemantikai összefüggés a másik fontos tényező a rekonstruálásnál. Azonos szintű szemantikai kategóriák esetén a rekonstrukció általában lehetetlen. Ez áll a többi határozói rag esetére is. Vö. például:

Veszek kenyér kolbász

esetét, amely lehet „Veszek kenyérhez kolbászt” vagy „Veszek kenyeret kolbászhoz” stb.

Összefoglalóan tehát azt mondhatjuk, hogy a grammatikailag helytelen mondatot csak akkor tudjuk a beszéd szituációtól függetlenül megérteni, ha egyértelműen rekonstruálható. Utóbbinak grammatikai és szemantikai feltételei vannak. A grammatikai feltételek általában megfogalmazhatók, ha megadjuk a szó-alakok generálásához szükséges modellt. A modell alapján megállapíthatók a különböző morféma-osztályok, amelyek a grammatikalitás fokának megállapításánál is szerepet játszhatnak. A szemantikai összefüggések tulajdonképpen szelekciós szabályok, megadják, hogy az ige milyen alany-nyal, milyen tárggyal, milyen kettős tárggyal stb. állhat. Ha valamely szó-

lánc eleget tesz, még hozzá egyértelműen tesz eleget, egy szelekciós szabálynak, akkor a szó-lánc minden esetben egyértelműen rekonstruálható.

Láttuk azt is, hogy a grammatikailag helytelen mondatoknak csak viszonylag kis hányada rekonstruálható. A többi legfeljebb a beszéd-szituáció ismeretében érthető meg.

Szerintünk a grammatikalitás fokait ilyen irányban is meg kellene állapítani. Itt természetesen csak a problémát vázolhattuk és csak részeredményeket közölhattünk.

Az ilyen természetű vizsgálódásoknak több szempontból van jelentősége. Bizonyos szempontból fényt vetne a nyelvtanulás folyamatának néhány eddig még tisztázatlan kérdésére, elsősorban ami a gyermekeknél az anyanyelv elsajátítását illeti. Másrészt jobban megértjük e tények világánál azt a körülményt is, hogy a grammatikailag helytelen mondatoknak tekintélyes részét megértjük. Szerintünk ezek a vizsgálódások azt is el fogják dönteni, hogy mikor milyen esetekben lehet formálisan eldönteni az érthetőséget és mikor következménye a beszéd-szituációnak (az első kérdés eldöntése maga után vonja a második eldöntését is). Végül, de nem utolsósorban a gépi szintézis néhány problémája is kapcsolódik e témakörhöz.⁸

KIEFER FERENC

⁸Itt elsősorban arra gondolunk, hogy azok, akik mondatok generálására írtak programot, különböző módszerekkel kísérleteznek a hibák kijavítására. A hibák kijavításánál a fenti megfontolások is szerepet játszhatnak.