

Традиция и нови теории в българското езикознание

Елена Тодорова (София)

Трудът *Формална граматика на българския език* (Осенова, Симов 2007) е новаторски в българското езикознание, тъй като се опира на синтактично анализиран корпус от реални текстове и следва съвременна формална лингвистична теория в описанията си.

Книгата е предназначена за „всички, които се интересуват от принципите на моделиране на езика и които не се страхуват да търсят нови предизвикателства в езиковите интерпретации“ (с. 6). Главната нейна цел е да разшири постигнатото досега в областта на „автоматичния анализ на българските изречения“, от една страна, и на „построяването на софтуерни анализиращи програми“ – от друга (с. 113).

Така е представена тя от авторите – доц. д-р Кирил Симов, член на Секцията за лингвистично моделиране в Института за паралелна обработка на информацията (ИПОИ) на БАН, сега Институт за информационни и комуникационни технологии (ИИКТ), и ръководител на нейни международни проекти, най-популярен между които VulTreeBank, и доц. д-р Петя Осенова, преподавател в Катедрата по български език на Софийския университет „Св. Климент Охридски“, със завършена математическа гимназия и специалността славянска филология в СУ. И двамата автори имат множество публикации по темата в България и в чужбина, част от тях са изнесени в списъка на използваната литература. Теоретичната основа на анализите им са принципите на така наречената *Опорна фразова граматика* (ОФГ), позната в чуждестранната литература под названието *Head-driven Phrase Structure Grammar* (HPSG).

В най-общ вид, като се придържаме към изложеното съдържание (3–4), можем да представим основните части на работата в последователността: Увод (5–6); изложение на формалните основи и универсалните принципи на Опорната фразова граматика (7–21; самостоятелна разработка на К. Симов) и на формалния ѝ модел в графичен вид (22–29); анализ на типовете лексикални елементи, покриващи десетте части на речта, плюс причастията и деепричастията (30–49) и на определени фрази (с подразделения, съсредоточени върху опората и зависимата част) – глаголна (50–70), именна (71–80; самостоятелна разработка на П. Осенова); фрази

на прилагателното и наречието и предложни (81–90); в отделна глава са разгледани координационните (91–95) и прагматичните фрази (103–110), както и някои елиптични конструкции (96–102) и кореферентни отношения (111–112). Следват кратко заключение (с. 113); преглед на използваната литература (114–115) и две приложения – списъци на синтактичните и на морфологично-синтактичните етикети, съдържащи се в проекта VulTreeBank (116–117 и 118–124).

В предлаганата дискусия вниманието е съсредоточено върху подбрани въпроси от по-общ характер, възникнали в процеса на запознаването с текста на работата. За по-точното им и пълно формулиране са привлечени допълнително и публикувани вече материали, свързани с разработваната тематика. Те са събрани в приложената тук библиография. По-долу следват компетентните отговори на самите автори, обобщени от П. Осенова и съобразени с интересите и на една по-широка читателска аудитория.

Е. Тодорова: Според така изложената от вас цел на работата, основен неин обект са българските изречения, т. е. тя е посветена на въпроси преди всичко от синтаксиса на българската граматика. Заглавието на монографията обаче, представено като *Формална граматика на българския език*, съдържа известно противоречие за обхвата на предстоящия граматичен анализ. Неясно е освен това за кой български език става дума – съвременен, книжовен и пр. Бихте ли направили – от гледна точка на сегашния ви прочит – съответните доуточнения или приемате, че това са пропуски, допуснати неслучайно?

П. Осенова: Всъщност книгата включва не само синтаксис, но също така и информация за частите на речта, за поведението на текстовите единици в дискурса. Във всяка граматика равнищата си взаимодействат. В този смисъл наистина понякога заглавията отразяват отделни идеи на авторите. Така например книгата на Р. Ницолова *Българска граматика. Морфология* (Ницолова 2008) всъщност включва цялата граматика, макар и през призмата на морфологията. Аз лично смятам подобен подход за предимство. В нашата книга основната призма е синтаксисът, но синтактичните явления не могат да бъдат обяснени без препратки към морфологията, лексикалното знание и дискурса. Това, което не позволи да опишем в детайли синтактичните и морфологичните особености на езиковите явления, беше ограничението в обема на изданието и кратките срокове за подготовката му. Но имаме идеята да преиздадем книгата, написана по-подробно и в по-достъпен формат.

За кой български език става дума? Става дума за съвременния книжовен български език, тъй като в корпуса от синтактични описания има текстове от съвременната българска литература, от съвременни ежедневници и от съвременни административни документи.

Е. Тодорова: Предварителният поглед върху съдържанието на работата разкрива известна неравнопоставеност (до прекалено неравностойно запълване) между отделните нейни части – отнася се преди всичко за главите от V до XIII, посветени на видовете фрази и представящите ги типове. Към това впечатление се прибавя и фактът, че на някои от тях (именните фрази) е посветена отделна монография (Осенова 2009), съдържаща освен прегледни и обширни въвеждащи части и впечатляващ по обхват илюстративен материал, извлечен от посочения

като използван и тук корпус (BulTreeBank)¹. Означава ли това, че и на останалите раздели – или на някои от тях (кои и защо) – ще бъде отделено подобно внимание? Освен това – интересно би било за читателя да бъде запознат и с принципите, определили включването (и изключването) на частите на материала и връзката му с последователността, в която той е изложен в традиционната граматика.

П. Осенова: Както споменах, една от причините беше ограничението в обема и кратките срокове, наложени от програмата-спонсор на изданието. Книгата „Именните фрази в българския език“ беше издадена две години по-късно (Осенова 2009). В новия вариант на формалната граматика смятаме да отделим балансирано внимание върху всички глави, които представят езиковите явления в ОФГ, тъй като наистина последните части са конспективни и само набелязват проблемите. Всъщност тази книга има за основа нашия отчет по проекта BulTreeBank (2001–2004; 2005–2007), спонсориран от Фондацията Фолксваген, Германия. Отчетът е написан на английски език, тъй като беше предназначен за фондацията. В него сме включили по-скоро отклоненията от традиционната граматика, защото те бяха новите и интересните явления за нас. После обаче осъзнахме, че едно описание трябва да включва и типичните случаи. Затова в българския вариант се постаряхме да включим и тях.

Освен това, в синхрон с традиционната граматика, последователно са представени частите на речта и след това – типовете словосъчетания. Новите аспекти идват от няколко посоки, като например от използването на съвременна езикова теория (което налага представяне на формализма и принципите ѝ); от специалното внимание към някои езикови явления като проблемите на координацията (съчинителното свързване), кореференцията, елипсата, дистантно разположените конституенти. Така че, от една страна, се предлагат нови, нетрадиционни решения, които са преосмислени при работата с големи масиви от текстове, а от друга страна, се дискутират явления, които често остават извън обхвата на традиционните граматика.

Е. Тодорова: В помощ – и от значение – за цялостното възприемане и осмисляне на текста би било, ако в работата – и по-точно в нейните начални части – се вмъкнат конкретни сведения и коментари за мястото на ОФГ в развитието на структурализма в езикознанието, от една страна, и за нейното съвременно използване в чужбина (и особено в България) – от друга. Във връзка с това бихте ли ни запознали по-подробно и с някои от темите и резултатите от работата по тях, осъществени в проекти между български и чуждестранни специалисти, за което споменавате в началото на книгата?

П. Осенова: Опорната фразова граматика е модерна лингвистична теория за моделиране на езика. Тя е от теориите, които разчитат на лексикалната информация в речника и на унификацията като метод на проекция на тази информация. Създадена е в САЩ от Карл Полард и Айван Саг през 1987 г., като следващото, преработено издание, е през 1994 г. Както всички модерни теории, и тя се опира на изначалните идеи на Ноам Чомски, но и се различава от тях. Напри-

¹ Срв. рецензията за тази монография – Тишева 2011 (бел. ред.).

мер не се работи с празни елементи и премествания; взема се предвид взаимодействието между всички лингвистични равнища, без да се абсолютизира ролята на синтаксиса. Друго предимство е това, че се отчита влиянието не само на фразите, но и на зависимостите между думите. Във връзка с това трябва да отбележа, че ОФГ комбинира конституентния (основан на фразите) и депendentния (основан на думите) тип анализ. Интересен е въпросът за ролята на ОФГ в развитието на структурализма в езикознанието. Ако структурализмът се разбира и като вид формална граматика, то ОФГ е теория, зад която има ясен граматически формализъм. Това означава, че може да се оценява колко добър е моделът на граматиката, кои явления са добре анализирани и при кои има проблем. За съжаление, в България ОФГ не е разпространена, тъй като изисква обем от знания, които не се предлагат от филологическите бакалавърски програми. В момента аз водя магистърски курс „Формални граматики“ в програмата „Лингвистика“ към Факултета по славянски филологии (СУ „Св. Климент Охридски“), което е една стъпка по посока на запознаване на българските студенти със съвременни лингвистични теории и приложението им за българския език. Освен това в ОФГ разработки има и доц. д-р Цветомира Венкова от Факултета по класически и нови филологии. Има и други българи, които работят в посочената теория, но главно в чужбина (напр. Таня Августинова от Университета в Саарланд). Теорията, освен за моделиране на естествените езици, се използва успешно както за целите на машинния превод, така и за автоматичен синтактичен анализ. Радвам се, че вече има и първа версия на ОФГ граматика за българския език (Осенова 2010), която е имплементирана и се използва в проект за машинен превод. Първата ѝ версия разработих в тясно сътрудничество с Дан Фликиндръ в Станфордския университет, където бях благодарение на стипендия Фулбрайт. Машинният превод е област, която комбинира усилията на лингвисти, компютърни лингвисти и информатици от доста време насам. Нашата задача в проекта е да направим ОФГ езиков модел на българския език, който да подобри статистическия превод от български език на английски език. Експериментите се оценяват количествено и качествено. И двата типа най-общо показват, че използването на лингвистична информация в подходящата комбинация от характеристики дава много добри резултати.

Някои колеги смятат, че формалните граматиките не работят със семантика, което е грешно твърдение. Напротив, ОФГ например работи със семантично представяне, наречено Семантика на минималната рекурсия (Minimal Recursion Semantics). Това представяне се използва и в други теории.

Е. Тодорова: В отделни глави на работата се тълкува употребяваната в тях терминология. Не би ли било по-ясно и прегледно за читателя, ако тя – или поне основната нейна част – се прибави като трети, самостоятелен списък в Приложението? Освен това в прекалено сложния език, с който се предава на места терминологичната и теоретичната информация в книгата, се забелязва известно отстъпление от достъпността му в предишни ваши изследвания (напр. в Осенова, Симов 2005). Нали на нея все пак би трябвало да се гледа и като на учебник, което намираме като образец в съвременните крупни синтактични изследвания на проф. Йордан Пенчев (Пенчев 1984; 1998)?

П. Осенова: Бих искала да отбележа, че книгите на Йордан Пенчев са също толкова сложни, колкото и нашата, и то по същите причини – или заради факта, че се изисква предварителна лингвистична подготовка от страна на читателите, или поради това, че написаното е по-формално. Според мен университетското ни образование следва все още до голяма степен традиционните параметри в езикознанието. В това няма нищо лошо, разбира се, но без началните познания по формална лингвистика няма как напълно да се разбере подобна литература. Неслучайно книгата ни се харесва и оценява предимно от колеги българи, които работят в чужбина. Но смятаме, че нашите студенти заслужават да бъдат обучавани в най-новите тенденции в лингвистиката – затова и полагаме усилия в тази посока.

Терминологията в преобладаващата част на текста е типичната терминология на синтаксиса като езиковедска дисциплина, но съм съгласна, че е по-добре да има индекс с термините за удобство.

Е. Тодорова: Един конкретен въпрос, свързан със зададените по-горе и засягащ значението и употребата на основния термин в труда. Това е терминът *фраза*, който тук не е специално дефиниран, а в други ваши работи (напр. в Осенова 2009: 9) е определен като синоним на *словосъчетание* и *група*. Едно поточно разяснение в текста на книгата, а тук – като отговор – може би не биха били излишни?

П. Осенова: Фразата е наистина словосъчетание, т. е. група думи, които заемат една синтактична позиция. Тук проблемът е по-скоро терминологичен – дали подобна структура да се нарече *група*, *фраза* или *словосъчетание*. Според мен се сблъскват различни езиковедски традиции и затова колегите-лингвисти имат различни предпочитания.

Е. Тодорова: Няколко въпроса, отнасящи се до перспективите, предоставяни от корпуса VulTreeBank (www.bultreebank.org) и подпомагащи – преди всичко – приложението му в областта на лингвистиката. Те са набелязани от вас в заключението на книгата, но конкретното им разясняване тук би привлякло вниманието на широк кръг специалисти, работещи в различни области на българското езикознание.

П. Осенова: Корпусите не са нещо ново – те съществуват от доста време². Откакто се развиха обаче информационните технологии, широко разпространение придобиха електронните корпуси. Ако говорим за синтактичния корпус, неговата роля е предимно за обучение на автоматични синтактични анализатори (парсери), за по-прецизно извличане на информация в езиковедски изследвания, машинен превод и др. Разбира се, подобен корпус може да се използва и за целите на обучението по българска граматика.

Е. Тодорова: Най-напред – бихте ли разяснили приложението на този специален текстов масив в лексикографската практика и по-точно – в изработването на различни видове речници (какви и как)? Ако такива вече съществуват, посочете ни техните заглавия, като ни ориентирате, макар и накратко, в процеса и особеностите на създаването им.

² Срв. напр. Лингорска 1978 (бел. ред.).

П. Осенова: За лексикографията по-ценни са по-големите корпуси. Такъв например е Българският национален референтен корпус – VulTreeBank с над 400 млн. думи. Подобни корпуси се използват за установяването на честотата на употреба на различните лексеми при изработване на речника. Така сме направили *Правописен речник на българския език* (Попов и др. 2003). Наскоро беше пусната и свободно достъпната уебслужба за търсене на думи и изрази в контекст. Така колегите могат да извличат най-разнообразни примери във връзка с конкретните си задачи. Ето и връзката: www.webclark.org. Нашите амбиции са да разширяваме тази услуга по посока на повече и лингвистично анализирани ресурси, от една страна, и по посока на различни обработки на текстове (сегментирани, морфологичен и синтактичен анализ) – от друга.

Корпусите също така са ценни и за извличане на актуални примери, нови значения и устойчиви съчетания или колокации, които да отразяват съвременната езикова ситуация в различни жанрове и регистри.

Е. Тодорова: Извънредно интригуваща е изтъкнатата от вас перспектива за помощта, която корпуси от този тип биха имали в преводаческата работа – ангажираща, както е известно, професионалисти с разностранни филологически интереси. Как по-точно те биха подпомогнали преводаческия процес?

П. Осенова: Така наречените паралелни синтактични корпуси подпомагат обучението на програми за автоматичен машинен превод. Това са корпуси между два или повече езика, които са подравнени на ниво изречение и/или дума и фраза. В идеалния случай текстовете и на двата (или повече) езика са анализирани синтактично. Разбира се, синтаксисът не е достатъчен за адекватен превод, необходима е и семантика. Както вече споменах, в момента работим по такъв проект (EuroMatrixPlus), като сме изградили паралелен корпус между текстове на български и английски език. Българската част е анализирана на морфологично, синтактично и семантично равнище.

Подпомагането на преводаческия процес може да стане по следния начин: преводачът има на разположение различни приемливи преводни варианти, между които може да избере или да се ориентира. Тези преводни варианти се натрупват и образуват т. нар. „преводна памет“. Освен това, като резултат от статистическия машинен превод, преводачите получават най-честите и най-типичните преводи на даден израз от един език на друг.

Е. Тодорова: Един въпрос, изискващ по-широк отговор, тъй като се отнася до ролята на синтактично аотираните текстове в обучението, т. е. както би трябвало да се подразбира – във важната и обществено значима сфера на образованието по български език. Тук разширявам този въпрос с отправка към сборника *Аргументна структура. Проблеми на простото и сложното изречение* (Коева 2005), в който е включена и вашата статия *Моделиране на лингвистичните явления в синтактично аотиран корпус* (Осенова, Симов 2005: 259–269). В предговора му се подчертава мисълта, че публикуваните в него съвременни синтактични изследвания представляват всъщност едно помагало, допълващо учебниците по синтаксис и лекционната работа на поколения преподаватели в СУ „Св. Климент Охридски“, ПУ „Паисий Хилендарски“ и Нов български университет. Прави впечатление твърдението, че изданието цели да подпомогне както

студентите от филологическите специалности на българските университети, така и работата на преподавателите – при това не само във висшите, но и в средните училища. Бихте ли коментирали съдържанието на този текст, като обхванете мястото на синтаксиса изобщо – и на съвременния синтаксис по-специално – в цялостното обучение по български език – във висшите и, особено, в средните училища: досега и в перспектива.

П. Осенова: Според мен, подобно помагало е подходящо най-вече за студентите и преподавателите във висшите училища, а не за учителите в средните училища. Помагалото включва както по-обзорни статии, така и конкретни изследвания – така че от преподавателя и студента зависи как ще се използват текстовете от сборника най-целесъобразно.

Ако се върна на нашата книга, то нейното място е по-скоро в обучението на магистранти и докторанти. Тя е на по-високо образователно равнище и предполага придобити основни познания по традиционен и формален синтаксис. Синтаксисът е важна дисциплина. Той се гради върху морфологията и е свързан със семантиката.

Смятам, че в училищата е добре да си остане традиционният, функционалният синтаксис. Но за университетското обучение е необходимо по-високо равнище на познанието.

Е. Годорова: Още една малка добавка към поставения по-горе проблем, възникваща във връзка с публикувани статистически данни и произтичащи от тях изводи за степента на типичност – и нетипичност – в употребата на различните типове именни фрази (вж. цитираната монография на Осенова 2009: 74–94). Очерталата се там картина – по принцип, независимо от някои непълноти в подхода към статистическото обработване на разнородния текстов материал – поставя въпроса и откъм обратната му страна. Става дума за възможността от използване на съвкупности от статистически лингвистични данни (не само синтактични), моделиращи и противопоставящи текстове с различна стилова принадлежност, в преподавателския процес. Как и къде – в по-широк план – би намерила мястото си според вас тази интересна и в теоретично, и в приложно отношение информация?

П. Осенова: Подобна информация би могла да намери мястото си на много места в учебния процес, стига преподавателят да е подготвен. Статистическите лингвистични данни биха помогнали и за установяване на тенденциите в употребата на българския език в различните видове дискурс. За чуждоезиковото обучение би било ценно най-напред да се изучават по-честите типове и/или явления, а след това – периферните.

Във връзка с това ще споделя, че ми се иска да виждаме повече корпусни изследвания от колегите-езиковеди, тъй като корпусните изследвания в същността си – независимо от статистическите непълноти – работят с големи масиви от данни и следователно показват много пълно и вярно езиковата действителност.

Е. Годорова: Може би бихте искали да добавите нещо, което аз съм пропуснала във въпросите си дотук?

П. Осенова: Да, бих искала да добавя следното. Това, което се случва, когато се работи с големи масиви от езикови данни, е че трябва да се създаде един общ

аналитичен модел за всички изречения в корпуса. Това означава, че когато се анализира дадено изречение, не можем да си позволим да го прескочим заради неясен механизъм за анализ. Обикновено в граматиките се дават типични примери, които не създават проблеми, но когато се анализират реалните текстове в тяхната пълнота и разнообразие, възникват и „неудобни“ случаи. В книгата си представяме модел, който работи за всички явления, но не твърдим, че решенията ни са най-правилните или най-елегантните. Има изключително голяма нужда от изучаване на синтаксиса на българския език във формален план. Нашата книга е един възможен поглед върху него, а не задължително предписание. Оказва се, че има явления, които не са проучени достатъчно (адвербиалните съчетания, ролята на фокусиращите думи в рамките на фразата и др.) и явления, които поставят трудности за всички езикови теории (като елипсата и координацията).

Е. Тодорова: И един последен въпрос, опиращ до пътищата на развитие, които стоят пред съвременната лингвистика, и по-точно – до коментар на съществуващата „конкуренция“ между формалния и хуманитарния подход към езиковия анализ. Т. е. между строгата и логически издържана типология – като негов резултат (ако не се стига до нежелан краен математически формализъм) и насочването на вниманието към говорещия субект – с изводи за текста като продукт на човека и обществото, към което той принадлежи, предоставящи материал и за социални, психически, културни, миросгледни, творчески и пр. интерпретации на факти и резултати от тяхното анализиране. Съществува ли възможност – от трета страна – за тяхното преплитане и взаимопроникване и каква е ролята на прагматиката и семантиката за това (последната – оставяна съзнателно от проф. Й. Пенчев като обект на предстоящо отделно и специално разглеждане (Пенчев 1984: 153; 1999: 655)? Бихте ли коментирали и най-характерното – в рамките на това противопоставяне – за езиковедската работа днес в по-широки граници – напр. в Германия (където К. Симов има съвместни проекти) и в САЩ (откъдето наскоро се завърнахте след специализация)?

П. Осенова: Не смятам, че има конкуренция между подходите (т. нар. формален и хуманитарен). Напротив, те са свързани и се допълват. Всяка лингвистична теория, която има за основа математически или логически формализъм, е добра теория, тъй като нейната точност и обхватност може да бъде проверявана. Разбира се, има представяне за ползване от човек и представяне за ползване от компютър. Тази разлика трябва да се отчита коректно, за да няма проблеми. Ние работим в лингвистичната теория ОФГ именно защото тя отчита ролята на всички равнища на езика – фонология, морфология, синтаксис, семантика, прагматика, информационна структура, и то в тяхното взаимодействие. Разбира се, кодирането на информация по подобен начин не е лесно и отнема време. Според мен обаче това е правилният път. Като добавим и адекватността на подобна теория за нуждите на компютърните приложения, предимствата да се работи в нея са много. Тази балансираност между теоретични постановки и реални приложения е изключително ценна.

Както споменах вече, аз не разглеждам развитието на лингвистиката като противопоставяне между математически и езиков подход. Чудесно е, че има български езиковеди като Мирослав Янакиев, Йордан Пенчев, Руселина Ницо-

лова, Елена Паскалева и други, които в една или друга степен в трудовете си виждат мост, а не конфликт между формалното и езиковото.

В работата си с колегите в чужбина също не усещам противопоставяне между формален и хуманитарен подход.

В предлаганата тук рецензия, или по-скоро дискусия, книгата на Петя Осенова и Кирил Симов беше използвана като повод за размишления и поставяне на някои принципни въпроси. Евентуалните нови рецензии, обхващащи и анализиращи нейното съдържание и от други страни, ще доведат до интересна и важна за развитието на българското езикознание дискусия, включваща въпроси и за това как би трябвало да изглежда една пълна и всеобхватна многотомна граматика на съвременния български книжовен език и за колективите, които биха я създали³.

Библиография

- Коева 2005: С. Коева (съст.) *Аргументна структура. Проблеми на простото и сложното изречение*. София: СЕМА РИШ, 2005.
- Лингорска 1978: Б. Лингорска. По някои проблеми на корпуса, предназначен за съпоставителни изследвания. – *Съпоставително езикознание*, 1978 (III), № 4, 3–14.
- Ницолова 2008: Р. Ницолова. *Българска граматика. Морфология*. София: УИ „Св. Климент Охридски“, 2008.
- Осенова 2009: П. Осенова. *Именните фрази в българския език*. София: ЕТО, 2009.
- Осенова 2010: P. Osenova. *Bulgarian Resource Grammar. Modeling Bulgarian in HPSG*. Saarbrücken: VDM, 2010.
- Осенова, Симов 2005: П. Осенова, К. Симов. Моделиране на лингвистичните явления в синтактично анотиран корпус. – В: С. Коева (съст.) *Аргументна структура. Проблеми на простото и сложното изречение*. София: СЕМА РИШ, 2005, 259–269.
- Осенова, Симов 2007: П. Осенова, К. Симов. *Формална граматика на българския език*. София: БАН, Институт за паралелна обработка на информацията, 2007.
- Пенчев 1984: Й. Пенчев. *Строеж на българското изречение*. София: Наука и изкуство, 1984.
- Пенчев 1998: Й. Пенчев. *Синтаксис*. – В: Т. Бояджиев, Ив. Куцаров, Й. Пенчев. *Съвременен български език. Фонетика. Лексикология. Словообразуване. Морфология. Синтаксис*. София: Петър Берон, 1998, 498–655.
- Попов и др. 2003: Д. Попов, К. Симов, С. Видинска, П. Осенова. *Правописен речник на българския език*. София: Наука и изкуство, 2003.
- Тишева, Осенова 2011: Й. Тишева, П. Осенова. *Именните фрази в българския език*. София: ЕТО, 2009. 187 с. – *Съпоставително езикознание*, 2011 (XXXVI), № 2, 99–103. [Рец.]

³ Книгата е свободно достъпна на следната връзка: <http://www.bultreebank.org/bgpapers/FormalGrammarBG.pdf>.