

***НАУЧНИ РАЗРАБОТКИ
НА ДОКТОРАНТИ ОТ ФАКУЛТЕТА
ПО СЛАВЯНСКИ ФИЛОЛОГИИ***

МОДЕЛИРАНЕ НА ЯВЛЕНИЕТО МОРФОЛОГИЧНА МНОГОЗНАЧНОСТ

Станислава Кънчева

Настоящата статия разглежда явлението *морфологична многозначност* и по-конкретно – моделирането му с оглед на теориите от типа **Дума и парадигма**, които са най-подходящи за описание на флективни езици, какъвто е българският. Моделирането се осъществява на базата на модела на Цвики (Цвики 1985), като за основа се използва формализмът му.

Ключови думи: морфологична многозначност, типове морфологична многозначност, моделиране, правила за експоненция, правила за съотнасяне

The paper discusses the morphological ambiguity and in particular its modelling in view of **Word and Paradigm** theory, which is the most suitable one for languages with inflectional word formation like Bulgarian. The author uses the model of Zwicky (Zwicky 1985) as basis of her representation.

Key words: morphological ambiguity, morphological ambiguity types, modelling, rules of exponence, rules of referral

Настоящата статия представя моделирането на явлението *морфологична многозначност*¹. Под морфологична многозначност се разбира свойството на лексикалните единици, съвпадащи в графичното си изписване, да принадлежат към различна част на речта или да имат различни граматически категории. Съответно обект на изследване са графемите, които могат да се отнесат към различни словоформи (независимо дали на една лексема, или на повече) и оттам – към различни граматически значения. За моделирането на типовете морфологични многозначности² е приложен моделът на А. Цвики (Цвики 1985). Като начало е нужно да се направят някои уточнения. Правилата, които се използват при Цвики, са реализационни, т.е. посредством тях се генерират форми. За целите на изследването на многозначностите обаче тези правила се използват основно

¹ За повече информация относно морфологичната многозначност вж. Кънчева 2015.

² Под *тип* се разбира *съвкупността от възможните граматически категории и граматически значения, които се носят от определени лексеми извън контекст*.

като формализъм³, според който да се представят единиците, участващи в определени типове морфологични многозначности, така че да може експлицитно да се наблюдават зависимостите между граматическите категории по типове. Това решение е обосновано и от факта, че според обема и нуждите на настоящото изследване не се строят цели парадигми, които да отговарят на лексикалните единици, срещащи се в типовете. Акцентът се поставя върху системните многозначности, т.е. тези с висока продуктивност. Целта е да се анализира морфологичната многозначност като системно явление и оттам да се проследят закономерности в появата на определени типове. За момента спорадичните срещания са извън обхвата на изследването, тъй като при тях няма закономерности и затова те трудно биха били моделирани. Поради тези причини се работи само на ниво основно правило, като няма да се използват правила за изключения и няма да се строят правила за целите парадигми, а само за редовете, за които е необходимо.

Друг момент, на който трябва да се обърне внимание, е въпросът за фонетичната форма. Реализационните правила от моделите от типа **Дума и парадигма** дават като резултат фонетична форма. Морфологичната многозначност се разглежда като отношение между графичната форма и граматическото значение, т.е. за анализа фонетичната форма не е от значение. За да се запази тази концепция, правилата няма да се строят с цел фонетичен облик и няма да се разглеждат фонетични вариации. Основно остава разделението *графична форма – граматическо значение*, което е посочено като ядро на явлението *морфологична многозначност*. Правилата имат план на израза и план на съдържанието, като план на израза е графичният облик, а план на съдържанието – граматическо значение. На този етап на анализ е прекъснатата връзката морфология–фонетика. На следващ етап на анализ може да се работи в две посоки – от една страна, към изграждане на връзка със синтаксиса и свързване на правилата с други правила, посредством които да се разрешават многозначности на основата на контекста; от друга страна, на базата на съществуващи правила се създават нови, чрез които може да се изгради фонетична форма, която да се използва за друг тип анализ. Според този анализ текст, илюстриращ устна реч, чрез правила за фонетичен облик може да доведе до графична форма и наличие на морфологична многозначност.

Моделът за описание на морфологични многозначности се придържа най-близко до модела на Цвики по отношение на формализма. Правилата имат съща-

³ Под *формализъм* в настоящата работа се разбира *формално представяне* в смисъла на математическата лингвистика.

та структура. Използват се квадратни скоби⁴, които да задават характеристиките. Това, което се намира между скобите, представлява една цялост. Скобите обаче не показват йерархията между граматическите категории⁵, а само конкретната им реализация. За анализ на морфологичните многозначности основно се използват частите на речта и граматическите значения, които притежават, както и конкретните граматическите характеристики, които са се проявили. Т.е. трябва да има отделни класове за част на речта, които да имат конкретна реализация (например *съществуващо име*), за граматически значения (лексикални, морфологични, лексикално-граматически), които също ще имат реализация (например *род, лице, вид*), и съответно грамемите, които са се реализирали при всяко значение (например *мъжки род, първо лице, свършен вид*). Има обаче и многозначности, при които деривацията оказва съществена роля за образуване на многозначността. Поради тази причина е необходимо да се предвиди и част в правилото, където да може да се реализира именно деривационен процес. Основната функция на правилата е да добавят флективен маркер, а не деривационен, и заради това суфиксите не могат да се оставят извън скобите и да се добавят заедно с тематични гласни и окончания. От друга страна, има случаи, при които деривационният елемент изиграва основна роля за поява на многозначността и е необходимо да бъде кодиран. На този етап на анализ най-доброто решение е да се съдържа самостоятелно – отделно от граматическите значения и от окончанията.

За описание на морфологичните многозначности се използват два типа правила – *правила за експоненция* и *правила за съотнасяне*⁶. Първо ще се спрем на правилата за експоненция. Тяхната функция е да представят връзката между графичната форма и граматическото значение. Граматическото значение се съдържа в две двойки квадратни скоби – в първата стои частта на речта заедно с тези нейни категории, които идват от речника или които не се изменят в парадигмата. Такива са лексикалните и лексикално-граматическите категории, но също и класификационните категории, каквото е спрежението. Във вторите скоби се намират категориите, които се променят в парадигмата, т.е. морфологичните. Там, където е налице и деривационен процес, който влияе за образуване на многозначността, се реализират трети скоби. В тях се съдържа деривационният маркер, който се появява като графична информация. След всички скоби се посочва

⁴ Подобен модел с квадратни скоби се използва и за описанието на именните фрази в българския език (Осенова 2009).

⁵ Подробна информация за йерархизирането на морфологичните категории може да се намери при Герджиков (Герджиков 2000), Алексова (Алексова 2008а; 2008б).

⁶ Примери и за двата вида правила са посочени по-надолу в изложението.

графичният облик на флективния маркер, който се появява. Флективният маркер се приема като основен елемент, тъй като той има най-голяма роля за образуването на системните морфологични многозначности.

При правилата за съотнасяне се използва същата организация със скоби, описана при правилата за експоненция. Разликата се състои в това, че с правилото за съотнасяне (правило 2) се цели да се посочи еднаквата реализация на две форми, като за основа се използва едно правило за експоненция за формата, приета като основна (правило 1), а второто правило означава еднаквата реализация на другата форма с тази на вече илюстрираната чрез първото правило. По този начин в правилото за съотнасяне не се посочва отново флективният маркер.

За по-голяма четивност на правилата избирам да работя с имената на частите на речта и граматическите им категории на български. Използват се следните съкращения, като са дадени по групи така, както се съдържат в правилата. Не са изброени всички части на речта и всички граматически значения, тъй като не се анализират всички типове многозначности, а само системните.

Списък на съкращенията, използвани в правилата за експоненция и съотнасяне:

ЧР – част на речта

СЪЩ – съществително име

ПРИЛ – прилагателно име

ГЛ – глагол

ПРИЧ – причастие

РОД – род със съответни възможни реализации – мъжки; женски; среден

ВИД – вид със съответни възможни реализации – св – свършен; несв – несвършен

ТРАНЗ – транзитивност със съответни възможни реализации – прех – преходен; непрех – непреходен

СПРЕЖ – спрежение със съответни възможни реализации – 1 – първо спрежение; 2 – второ спрежение; 3 – трето спрежение

КЛАС – за разграничение на причастията със съответни възможни реализации – мин_стр – минало страдателно причастие; мин_св – минало свършено деятелно причастие; мин_несв – минало несвършено деятелно причастие

ЧИСЛО – със съответни възможни реализации – ед – единствено число; множ – множествено число; бр_ф – бройна форма

ЛИЦЕ – със съответни възможни реализации – 1 – първо лице; 2 – второ лице; 3 – трето лице

ВРЕМЕ – със съответни възможни реализации – сег – сегашно време; аор – минало свършено време; имп – минало несвършено време

ОПРЕД – определеност със съответни възможни реализации – чл – членувана форма; чл_п – членувана форма с пълен член; чл_кр – членувана форма с кратък член; нечл – нечленувана форма

НАКЛ – наклонение с възможна реализация – пов – повелително

ДЕРИВ – за означаване на деривационен елемент, който оказва влияние за образуване на многозначността.

Тук като реализация се съдържа самият елемент или наставката, тъй като в настоящата работа се разглеждат само наставки.

Относно списъка с граматически категории е нужно да се направят някои уточнения. Първо, в настоящата работа се приема разпределението на граматическите категории, застъпено от Р. Ницолова (Ницолова 2008)⁷. При извличането на многозначностите се придържам към граматическите значения, приписани на думите според тагсета, приет от К. Симов и кол. (Симов и кол. 2004). Според тази позиция до момента например в граматическите значения на глаголите не се включва информация за спрежението. С оглед на последващия анализ обаче е нужно да се осъществи промяна и тя се състои в добавянето на категорията спрежение при глаголите⁸. Тъй като спрежението е информация, която идва от речника, тя ще се задава в първите скоби. Няма промяна по отношение на останалата информация от таговете.

Второ, относно разделението на категориите на такива, които идват от речника и не се променят в цялата парадигма, и такива, които се изменят, е нужно едно уточнение по отношение на рода – при съществителните имена родът попада в първите скоби и се разглежда като лексикална категория, докато при прилагателните имена и причастията родът преминава във вторите скоби, тъй като вече е категория, която се изменя в парадигмата.

Трето, при наклонението има само една възможна реализация – повелително наклонение, тъй като само тя оказва влияние при образуването на морфологична многозначност. За момента по подразбиране се реализира изявително наклонение, ако изрично не е посочено друго.

Четвърто, при категорията време се приемат само трите реализации – сегашно, аорист и имперфект, тъй като в таговете се съдържа информация само за тези времена и на това ниво на анализ се разглежда многозначност само между

⁷ И. Куцаров представя друга гледна точка (Куцаров 2007).

⁸ По-надолу в текста стават ясни причините.

синтетичните форми за времената. Аналитичните времена не са обект на изследване в настоящата работа⁹.

Имената на категориите се означават с главни букви, а конкретните им реализации – с малки. Връзката между категорията и реализацията е двуетоchie. Изреждането на категориите става чрез запетая. В правилата за експоненция могат да се появят три двойки квадратни скоби – в първата се описва частта на речта и лексикални или лексикално-граматически категории, които притежава – при съществителното име – род, при глагола – вид¹⁰. В тези скоби също се посочва преходността и спрежението при глагола. Преходността има значение от синтактична гледна точка или, както посочва И. Куцаров, глаголите се делят на преходни и непреходни според „някои лексико-синтактични особености“ (Куцаров 2007: 106). Спрежението е формален клас, който се изразява чрез завършека на сегашната основа. И преходността, и спрежението се посочват в първите скоби заедно с вида, тъй като те са характеристики, които принадлежат на лексемата и не се изменят в парадигмата. При причастието в първите скоби се съдържа информация за конкретното причастие, което участва в типа (например минало страдателно, минало свършено деятелно и др.), също и информация за вида и преходността на глагола, от който е образувано. При прилагателните имена в първите скоби има информация само за частта на речта. Във вторите скоби се показва реализацията на всички категории, които се изменят в парадигмата: при съществителното – число, определеност; при глагола – време, лице, число, наклонение (само където оказва влияние за появата на многозначността); при причастието – род, число, определеност; при прилагателните – род, число, определеност. Третите скоби съдържат информация за словообразователна морфема, ако тя има отношение към образуване на типа многозначност.

Едно правило за експоненция би изглеждало по следния начин:

Комбинацията [ЧР: СЪЩ, РОД: мъжки], [ЧИСЛО: ед, ОПРЕД: чл_кр] се реализира чрез **-а**.

Едно правило за съотнасяне би изглеждало по следния начин:

Комбинацията [ЧР: ГЛ, ВИД: несв, ТРАНЗ: прех, СПР: 2], [ВРЕМЕ: сег, ЛИЦЕ: 3, ЧИСЛО: ед. ч.] има същата реализация като комбинацията [ЧР: ГЛ, ВИД: несв, ТРАНЗ: прех, СПР: 2], [ВРЕМЕ: аор, ЛИЦЕ: 2, ЧИСЛО: ед. ч.] и като

⁹ Това решение е разгледано подробно в Кънчева 2015.

¹⁰ В настоящата работа се придържахме към становището, че видът е класифицираща, лексикално-граматическа категория, което се подкрепя от Л. Андрейчин (Андрейчин 1944), Ст. Стоянов (Стоянов 1980), И. Куцаров (Куцаров 2007), Р. Ницолова (Ницолова 2008).

комбинацията [ЧР: ГЛ, ВИД: несв, ТРАНЗ: прех, СПР: 2], [ВРЕМЕ: аор, ЛИЦЕ: 3, ЧИСЛО: ед. ч.].

Изграденият модел се прилага според разделяне на многозначностите в зависимост от броя части на речта, които участват в типовете морфологични многозначности. Акцент се поставя върху системните многозначности. Разделянето на системните типове морфологични многозначности може да се представи по следния начин:

1. Многозначност вследствие на съвпадения на флективни морфеми, т.е. на многозначност на флективни морфеми – тук основно участие вземат съществителните имена и глаголите, като най-често срещаните многозначни морфеми са *-и* и *-а*.
2. Многозначност като пресечна точка на деривация и флексия – съвпадението е резултат от пресичане на словообразуване и формообразуване, като тук основна роля имат причастията.
3. Многозначност, при която съвпаденията са следствие на определени процеси като конверсия и субстантивация – в този случай основно участват прилагателни имена, наречия, съществителни имена. В голяма част от примерите, които се наблюдават, се разглежда многозначност на ниво лексеми.

Преди да се представи моделирането на типовете, е нужно да се уточни дали и трите групи ще се описват чрез създадения формализъм. Основното предимство на правилата за експоненция е, че представят връзката между граматическото значение, като кодират граматическите категории и конкретните им стойности, и формалното му изразяване чрез граматическа морфема. По този начин формализмът покрива точно многозначностите, които попадат в първата група. При тях в най-голяма степен може да се наблюдават и зависимости между граматическите значения, които създават предпоставки за появата на многозначността. Добавянето на трети скоби с елемент ДЕРИВ (= деривация), който позволява кодирането на информацията относно словообразуването, решава въпроса за представянето на втората група многозначности. Предизвикателство обаче остава третата група, в която многозначностите много често са на ниво лексеми или в други случаи – в основата стои процес. Например при един от най-често срещаните типове между *прилагателно*, *среден род*, *единствено число*, *нечленувано* и *наречие за начин* в основата стои конверсията. Тук многозначността не може да се опише чрез ролята на флективни или деривационни елементи, нито появата ѝ може да се сведе само до графичното съвпадение на две

форми. Този тип е показателен за многозначности, които трябва да се кодират на ниво речник – потенциално всяка форма на прилагателно име за *среден род*, *единствено число*, *нечленувано* може да премине в класа на наречията. Затова позицията, която се приема на този етап на анализ, е многозначностите вследствие на процес да не се моделират.

С оглед на разглеждането на лексеми и ролята на лексикона е нужен по-специален поглед и към още една група многозначности. Става въпрос за типове, при които влияние оказват категории, идващи от речника. В тази група попадат многозначностите между двувидовите глаголи, оттам и породени от същата причина многозначности между причастия, между съществителни с разлика в рода. В тези случаи многозначността е основно на ниво лексема и оттам при някои типове преминава във всички форми (например при двувидовите глаголи). По същия начин, както и в описания по-горе случай, тук основно значение има информация, която се кодира в речника. Да вземем по-конкретно многозначностите при причастията. В тази група високата продуктивност на типовете с най-много представители се дължи основно на две причини – съвпадение на причастията на двувидови глаголи и съвпадение на имперфектното и аористното причастие на глаголите от трето спрежение и втори и трети разред на второ спрежение. Многозначността тук е между парадигми. Разлика има в стойностите на категориите от речника. В случаи като този основна роля има информацията в лексикона и в повечето случаи цялата лексема. За да се представят тези многозначности, е необходимо допълнително изследване и анализ, които надхвърлят обема и целите на настоящата работа. Поради тази причина тези многозначности няма да бъдат моделирани.

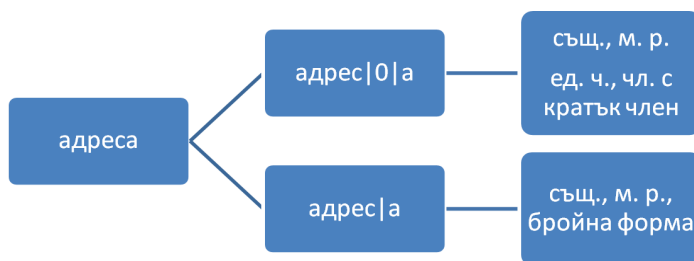
Тук е мястото да се спрем и на едно важно решение относно моделирането на многозначности. Става дума за флективността и въпроса колко позиции се приема, че се запълват от граматически морфеми. Р. Ницолова посочва, че думите във флективните езици имат само едно окончание, но в български това се нарушава, ако се приеме, че членуваните форми на съществителните имена имат две окончания (Ницолова 2008: 13). От друга страна, тематичната гласна заема отделна позиция. Настоящото изследване не си поставя за цел да решава въпроса колко точно позиции се реализират във флективните езици. Поради тази причина не се придържахме към традиционното представяне чрез позиции, тъй като основната цел е да се представят морфологичните многозначности, а формалните причини за тяхната поява са различни. Поради този факт се приема, че правилата попълват една позиция, като в нея може да застава всякаква граматическа

морфема (тематична гласна, окончание, определителен член). За да се запази коректното лингвистично представяне, където има повече от една граматическа морфема (в това число и нулева морфема), се означават и двете чрез знак +, който показва едновременна реализация. Прието е да се означава и нулевото окончание, което е налице в българската морфология и се приема от Ницолова (Ницолова 2008). В настоящата работа се приема становището, че има нулево окончание и то се кодира в правилата като цифрата нула – 0.

Илюстрираме разгледания формализъм чрез най-често срещания тип в рамките на съществителните имена (който е и един от най-продуктивните въобще), а именно между **[съществително нарицателно, мъжки род единствено число, членувано с кратък член]** и **[съществително нарицателно, мъжки род, бройна форма]**. Той се представя чрез следните правила:

1. Комбинацията [ЧР: СЪЩ, РОД: мъжки], [ЧИСЛО: ед, ОПРЕД: чл_кр] се реализира чрез **0 + -а**.
2. Комбинацията [ЧР: СЪЩ, РОД: мъжки], [ЧИСЛО: бр_ф] се реализира чрез **-а**.

При моделирането на този тип многозначност и двете правила са за експоненция. Формално графемите съвпадат, но пътят, по който са съвпаднали, е различен. Въпреки че са от една парадигма, не могат да бъдат представени чрез правило за експоненция и правило за съотнасяне, тъй като в настоящата работа е прието да се изразява и нулевото окончание, което означава, че реализацията на двете форми не е еднаква. По-добре това може да се види на Фигура 1.



Фигура 1

Както се вижда на схемата, графемата *адреса* отвежда към две словоформи. Едната е *адреса* с граматическо значение *съществително нарицателно, мъжки род единствено число, членувано с кратък член*, а другата – *адреса* с граматическо значение *съществително нарицателно, мъжки род, бройна форма*.

ма. Многозначността тук е породена от формалното съвпадение на краткия определителен член за мъжки род и окончанието за бройната форма. На схемата също така се вижда и формалната разлика, тъй като при първата форма е налице нулево окончание за мъжки род и след него се помества определителният член, докато при втората форма директно се поставя окончанието за бройна форма, което носи и значението мъжки род. Така че тук може да се разглежда формално съвпадение на определителния член и окончанието за бройна форма, но при детайлното представяне на двете словоформи може да се наблюдава дори и чисто графично разликата.

Представеният в настоящата статия формализъм илюстрира закономерностите между добавянето на граматически морфемии и граматическите значения, които те изразяват, и влиянието им за образуване на морфологични многозначности. Правилата за експоненция и съотнасяне са използвани за описание на две групи системни многозначности – тези, при които влияние оказват граматични морфемии, и тези, при които се наблюдава пресичане на формообразуване и словообразуване. Една част от многозначностите (основно тези, при които е налице протичането на процес) е извън обхвата на моделирането, тъй като е необходимо допълнително проучване, което надхвърля обема на настоящото изследване. Като основен извод от моделирането на многозначностите може да се изведе фактът, че начинът на функциониране на правилата ясно показва фузионните характеристики на българския език. Важно е да се отбележи, че правилата могат да се използват като основа за изграждане на компютърен модел. Подобен модел може да се приспособи за нуждите на компютърната обработка на естествен език. Също така е от значение и фактът, че правилата са придружени от фигури, илюстриращи типа многозначност с конкретен пример. Този подход би могъл да се използва при обучението по български език като роден и като чужд.

Библиография

- Алексова 2008а*: Алексова, К. За междукатегориалните връзки на морфологичното равнище. – In: *Bulgaristica-Studia et Argumenta. Festschrift fuer Ruselina Nitsolova zum 65. Geburtstag.* Herausgegeben von Sigrun Comati. *Specimina Philologiae slavicae.* Band 151. Begruendet von Olexa Horbatsch und Gerd Freidhof. Herausgegeben von Peter Kosta, Holger Kusse und Franz Schindler. Muenchen: Verlag Otto Sagner, 2008, 76–88.
- Алексова 2008б*: Алексова, К. Видове междукатегориални връзки в парадигмата на българския глагол (начални наблюдения). – В: *Матрицата – властта на подобие.* София: УИ „Св. Климент Охридски“, 2008, 230–247.

- Андрейчин 1944*: Андрейчин, Л. Основна българска граматика. София: Хемус, 1944.
- Герджиков 2000*: Герджиков, Г. Организацията на многочленните морфологични парадигми. – Език и литература, 1/2000, 150–180.
- Куцаров 2007*: Куцаров, И. Теоретична граматика на българския език. Морфология. Пловдив: Паисий Хилендарски, 2007.
- Кънчева 2015*: Кънчева, С. Морфологичната многозначност в българския език. Дисертационен труд, 2015.
- Ницолова 2008*: Ницолова, Р. Българска граматика. Морфология. София: УИ „Св. Климент Охридски“, 2008.
- Осенова 2009*: Осенова, П. Именните фрази в българския език. София: ЕТО, 2009.
- Симов и кол. 2004*: Simov, K., P. Osenova, M. Slavcheva. ВТВ–TR03: BulTreeBank Morphosyntactic Tagset. BulTreeBank Project Technical Report No. 03. 2004.
- Стоянов 1980*: Стоянов, Ст. Граматика на българския книжовен език. 3. издание. София: Наука и изкуство, 1980.
- Цвики 1985*: Zwicky, A. M. How to describe inflection. – In: Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistic Society, Mary Niepokuj, Van Clay, Vassiliki Nikforidou and Deborah Feder (eds.). Berkeley: Linguistic Society, 1985, 372–386.