

ЗА ЕДИН СПЕЦИФИЧЕН ТИП АНАФОРИ В БЪЛГАРСКИЯ ЕЗИК. ПРИЛАГАНЕ НА СЪЩЕСТВУВАЩИ ТИПОЛОГИИ

Зара Кънчева
Софийски университет „Св. Климент Охридски“

ON A SPECIFIC TYPE OF ANAPHORA IN BULGARIAN. AN APPLICATION OF EXISTING TYPOLOGIES

Zara Kancheva
Sofia University “St. Kliment Ohridski”

The paper attempts to present an adaptation of the existing typology of associative anaphora. The chosen model is based on a scheme that has already been used in a previous annotation project and whose validity has been tested with examples from the contemporary Bulgarian press. The categorization follows the semantic approach and it includes seven types of relations: metareference, class, temporal relations, dissimilation, entailment and attribution.

Key words: associative anaphora, coreference

I. Въведение

Една от основните характеристики на текста е неговата кохерентност, чийто основен инструмент е референцията – съотнасянето на езиков знак с екстралингвистичен факт. Поради същността на текста по-значимо е не единичното, а многократното рефериране на един и същи извънезиков обект чрез няколко различни номинации, което се нарича *корелация*, според определението на Е. Добрева и И. Савова (Добрева и Савова/Dobрева and Savova 2000: 185).

Друг подобен на корелацията езиков елемент, който също е градивна част на кохерентността на текста, е анафората. Анафората е сравнително по-слабо разгледана в специализираната литература и често е погрешно отъждествявана с корелацията. Основната разлика между двете явления е необходимостта на анафората от предварително споменат в контекста антецедент, чрез който интерпретацията на анафората е възможна; обикновено анафората и антецедентът са съгласувани. Анафората е характерна основно за именната група, но може да засяга и събития,

изразени с глаголи. Дефиницията за *типична анафора* намира своя аналог в българското езиковзнание в определението на Е. Добрева и И. Савова за *анафорични конектори* (Добрева и Савова/Dobрева and Savova 2000: 210) – анафоричният конектор препраща възприемащия към antecedента си и най-често представя синонимен или местоименен израз.

Често анафората и кореференцията се появяват едновременно в текста, но това не е задължително (Зиканова/Zikanova и др. 2015: 31), както личи от примерите:

- (1) Джон има книга. Мери също има една.¹
- (2) Харесвам София. София е един от най-красивите европейски градове.

Обект на това изследване е един специфичен тип анафора, който до известна степен се изравнява с концепцията на Е. Добрева и И. Савова (Добрева и Савова/Dobрева and Savova 2000: 186) за *частична кореференция* – отношение на взаимна включеност между референтите (*врата – брава*), и в по-тесен план – с дефинициите на В. Радева за *хиперонимия* и *конверсия* (Радева/Radeva 2017: 168-172). Този тип анафора в различните разработки се среща като *свързваща анафора* (*bridging anaphora*) (Clark/Кларк 1975), *индиректна анафора* (*indirect anaphora*) (Schwarz-Friesel/Шварц-Фризел 2007), *асоциативна анафора* (*associative anaphora*) (Charolles, Mieville/Шарол, Мевил 1999) и *свързваща релация* (*bridging relation*) у редица други автори и най-общо бележи релация между некореферентни изрази.

Един от най-ясните примери за изследвания вид анафора е връзката част – цяло:

- (3) Купих си старинен гардероб. Вратите са развалени и скърцат.

Структурата на статията е следната: следващата част представя кратък преглед на съществуващите типологии на специфичните анафори; в третата се прилага избраната класификация върху български примери, а в последната част е представено заключението.

¹ Примери (1), (3), (11), (18), (20) и (22) са преведени от изследването на Зиканова (2015).

II. Съществуващи типологии на специфичната анафора

В лингвистиката има разнообразие от класификации на особения тип анафори, но общото между повечето от тях е, че следват семантичния принцип – типовете анафори са базирани на смисловите връзки между изразите. Много от типологиите включват група анафори с неопределен статут, озаглавявани обикновено „други“, „неопределени“, „разни“. В голям брой класификации са изведени категории, съдържащи релациите *клас – случай* и *част – цяло*, редом с темпорални и функционални групи.

Важно е да се отбележи, че типологиите на специфичната анафора до голяма степен съвпадат с класификацията на концептуално-семантичните отношения в българската лексикология (Колковска/Kolkovska 2005).

Първа по време на възникване е класификацията на Х. Кларк (Clark/Кларк 1975), който въвежда термина *bridging*. Авторът обособява седем типа специфични анафори – *елемент – множество* (*set membership*), *индиректна референция по асоциация* (*indirect reference by association*), *индиректна референция по характеристика* (*indirect reference by characterization*), *причина* (*reason*), *основание* (*cause*), *последствие* (*consequence*), *съдействие* (*concurrency*).

Други автори представят типология от шест типа – *синонимия/хипонимия/меронимия* (*synonymy/hyponymy/meronymy*), *имена* (*names*), *съставни съществителни* (*compound nouns*), *събития* (*events*), *тема на дискурса* (*discourse topic*), *извод* (*inference*) (Poesio, Vieira, Teufel/Пое시오, Виейра, Тойфел 1997).

Повече видове са определени в класификацията на Гарден и кол. (Gardent et al./Гарден и др. 2003) – *множество – подмножество* (*set – subset*), *множество – елемент* (*set – element*), *събитие – аргумент* (*event – argument*), *индивид – функция* (*individual – function*), *индивид – атрибут* (*individual – attribute*), *цяло – част* (*whole – part*), *цяло – парче* (*whole – piece*), *индивид – елемент* (*individual – stuff*), *множество – член* (*collection – member*), *място – площ* (*place – area*), *цяло – подчаст* (*whole – temporal.subpart*), *локация – обект* (*location – object*) и *време – обект* (*time – object*).

Класификациите, насочени основно към аотирането в областта на компютърната лингвистика, най-често имат между четири и шест основни типа и по един допълнителен, в който попадат сложни и по-различни случаи (GCBT: Greek Coreference & Bridging Team 2014).

Типологията, разглеждана в това изследване, е опит за компилиране и усъвършенстване на няколко известни съществуващи класификации и адаптирането им към анафорите в полския език (Ogrodnichuk/Огородничук и др. 2016: 16). Авторите първоначално извеждат седем типа с общо петнадесет подтипа, но след извършване на проверки установяват редица слабости на тази система (разтегнатост на типологията, неяснота в дефинирането на някои секции и основно – прекалено много неопределени типове, наричани *други, останали*).

Финалният вариант, който адаптират, е съставен отново от седем типа с шест подсекции: *метареференция*, *клас*, *дисимиляция* (*темпорална и контекстуална*), *структурална анафора* (*агрегация и композиция*), *следствие*, *аналогична анафора* (*прилика и контраст*) и *атрибуция*.

Важно е да се отбележи, че авторите, изработили тази класификация, налагат ограничение по отношение на глаголната група и не включват анализиране и прилагане на примери с процеси и събития.

Първият тип е *метареференция* – този тип съдържа релации по моделите *има – модел* (*has – model*), *има – име* (*has – name*) и *има – марка* (*has – label*). От други автори типът е определян като металингвистичен. Отбелязва се, че анафората и antecedентът не преpraщат към обект от действителността.

Втората група е *клас* и към нея са отнесени релациите *клас – отделен случай*. Тези анафори са близки до типа на структуралните, които разглеждат връзки като *част – цяло*. Тази категория присъства под различни наименования в почти всички съществуващи типологии. Тя съдържа отношения на *хиперонимия* и *хипонимия*.

Типът *дисимиляция* съдържа темпорални отношения. Стандартните темпорални изрази обикновено включват глаголи в състава си, така че не са обект на тази типология. Класически пример за категорията представлява конструкцията със събитие и местоимение – вж. пример (10).

Четвъртият тип от типологията е *структуралната анафора*. Тя е разделена на две подсекции – *агрегация* и *композиция*, които също имат разновидности. Това е типът, при който има най-малко разлики в подхода на лингвистите и който е най-непротиворечив. Той е дълбоко обвързан с понятието за меронимия (връзки *част – цяло*).

Аналог на тази категория е интересната разработка на Кларк (Clark/Кларк 1975), който дефинира *необходими/вероятни/индуцирани части* (примери (12), (13)).

Типът *последствие* съдържа *причинно-следствени релации* като например *война – окупация*, *дим – миризма*, *конкурс – резултат* и други. Ярки представители на този тип в българския език са редица взаимосвързани глаголи (*спя – хъркам*, *спя – сънувам*).

Аналогичната анафора е разделена на *прилика* и *контраст*. Първият подвид е лесно откриваем, защото обикновено е придружен от изразите „като“ и „както“ – вж. пример (20), а вторият – защото се състои от отношения между *съхипоними* (думи с общ хипероним) – вж. пример (21).

Последният тип от класификацията е *атрибуция* – този тип включва релации между даден обект и мнението на някого за него; изразява вярване, съмнение, несигурност. Авторите отбелязват, че в много случаи този тип анафори се определят като кореференция, но когато съществуват различни мнения в един дискурс, това определение е неподходящо (Ogrodnichuk et al. 2016: 18).

III. Прилагане на избраната типология към българския език

Специално адаптиране на полската типология не е необходимо, защото, както може да се установи от следващите примери, всеки изведен в нея тип се реализира и в българския език.

1. Метареференция

- (4) „Чудовище“ – Детето сложи пръста си на думата, на която спря да чете.²

² Примери (2), (4), (5), (6), (7), (8), (9), (15), (16), (17), (18), (19) и (23) са експерирани от интернет платформите на вестниците „Капитал“ (2018) и „24 часа“

Пример (4) ясно демонстрира вида референция и е по-често срещан в българския език, спрямо конструкции, подобни на тези от примери (5) и (6).

- (5) *Ходих в ресторант „Пицата на Джани“. Тяхната пица изобщо не ми хареса.*
- (6) *Младата актриса често е забелязвана да пие вино и да хапва любимите си спагети в столичния ресторант „Спагети къмпъни“.*

2. Клас

Както личи от примерите, типът *клас* е често срещан и лесно разпознаваем.

- (7) *Хорър филмите залагат основно на ужасяващи кадри. „Писък“, от своя страна, има много добър сценарий.*
- (8) *Черните мури са широко разпространени в българските гори, но Байкушевата мура в Пирин се отличава с внушителната си възраст – повече от 1300 години.*
- (9) *Любовните романи са считани за банални, но новият роман на писателката разбива стереотипите.*

В примерите (7), (8), (9) антецедентите (*хорър филмите, черните мури, любовните романи*) отбелязват класове, а анафорите са конкретни представители на този клас (*„Писък“, Мария, Байкушевата мура, новият роман на писателката*).

3. Дисимиляция

(10) *Договорът с американската фирма е сключен през февруари миналата година при голямата обезценка на лева. Тogaва Булгаргаз продаваше гориво с огромна загуба поради нереално ниските цени за вътрешния пазар и скока на долара.*

- (11) *Варшава е много красив град, но предвоенната Варшава е била още по-красива.*

За този тип е важно да се отбележи, че във всички типологии има изведена категория, разглеждаща темпоралните отношения, независимо как е именувана.

(2018), както и от българския национален референтен корпус – BulTreeBank, през системата webclark.org.

4. Структурална анафора

Агрегация – множество – подмножество (set – subset), множество – елемент (set – element):

(12) *Влязох в стаята. Подът е дървен.*³

(13) *Джон е убит вчера. Убиецът е избягал.*

В пример (13) се наблюдава концепцията на Кларк (Clark/Кларк 1975) за *необходими части* – в случая убиецът е предполагаема част на събитието „убийство“.

(14) *БДЖ се е допитала до четири компании за изчисляването на парите, необходими за купуване на новите моторни влакове [...] Компанията вече обяви, че за първите 42 броя моториси, които иска да купи до 2023 г., ще са ѝ необходими 450 млн. лв.*

Композиция: цяло – част, цяло – парче, цяло – субстанция:

(15) *Ресторант Сириус ще вдъхне морския аромат и романтиката за бала, смятат от администрацията на комплекса.*

(16) *Направих торта. Изядох едно парче.*

(17) *Фирмата започва да произвежда пътни чанти през 2003 г. Първоначално кожата е внос от Турция.*

Авторите на типологията отбелязват, че *структуралната анафора* е една от най-често срещаните категории в направения от тях анализ за полския език. От изреченията (12 – 17) е видно, че примери от този тип са характерни и за българския език.

5. Последствие

Множество примери за тази категория анафори, включващи глаголната група, могат да се извлекат лесно, но те не са обект на настоящото изследване.

В примери като (18) и (19) възприемащият без затруднение интерпретира естествената причинно-следствена връзка между *огън* (антецедент) и *дим* (анафора); и *конкурс* (антецедент) и *резултат* (анафора).

³ Примери (12) и (13) са преведени от изследването на Кларк (Clark/Кларк 1975).

- (18) Наводнението унищожи десетки сгради в района. Собствениците на оцелелите къщи сега се борят с влагата.
- (19) Специалистът по обработка на данни Сет Стивънс-Давидовиц получава достъп до анонимните търсения в „Гугъл“, „Фейсбук“, „Твитър“ [...] Резултатът е впечатляващ.

6. Аналогична анафора

- (20) Главата му е огромна. Като голям балон.

Връзката между антецедента и анафората се онагледява до голяма степен благодарение на характерния за сравнителни конструкции предлог *като*.

- (21) Някои хора предпочитат да отглеждат *куче*, а други – *котка*.

В пример (21) отношението на контраст е постигнато чрез релацията между съхипонимите *куче* и *котка*, с хипероним – *животно*.

7. Атрибуция

- (22) – Как се казва *съпругът на Мери*?
– Джон, предполагам.
– Аз пък мисля, че е *Питър*.

Както личи от посочения пример (22), *съпругът на Мери* може да не се казва нито *Джон*, нито *Питър*, следователно тази категория анафори не бива да се отъждествява с понятието за кореференция, защото представя различни аспекти на един обект или същия обект, погледнат от различни гледни точки – вж. пример (23).

- (23) Той разполага с огромен капитал, който би бил жизненоважен за оцеляването на *Tesla*. Първоначалните индикации са, че от Саудитска Арабия нямат интерес към една непублична *Tesla* [...].

IV. Заключение

Разликите в поведението на специфичния тип анафора в българския език в сравнение с други езици са малки – състоят се в честотата на поява на различните типове анафори. Най-висока честота на поява в речта бележат категориите *клас* и *структурал-*

на анафора – типовете, съдържащи отношения на хиперонимия, хипонимия и меронимия, а най-рядко се откриват примери, илюстриращи типовете *метареференция* и *атрибуция*.

Подобен тип изследвания допринасят за влагането на повече лексикални релации в синтактичното изразяване на анафоричните отношения, а от друга страна – полагат основи за изграждане на българска типология на специфичната анафора.

БИБЛИОГРАФИЯ

- Добрева, Савова/Dobрева, Savova 2000:** Добрева, Е., Савова. И. *Текстлингвистика*. Шумен: УИ „Епископ Константин Преславски“, 2000. [Dobрева, Е., Savova. I. *Tekstolingvistika*. Shumen: UI “Episkop Konstantin Preslavski”, 2000.]
- Zikanova/Зиканова 2015:** Šárka Zikánová, Eva Hajicová, Barbora Hladká, Pavlína Jínová, Jirí Mírovský, Anja Nedoluzhko, Lucie Poláková, Katerina Rysová, Magdaléna Rysová, and Jan Václ. *Discourse and Coherence. From the Sentence Structure to Relations in Text*. Prague, Czech Republic: Institute of Formal and Applied Linguistics, 2015.
- Clark/Кларк 1975:** Clark, H. *Bridging*. // Proceedings of the 1975 Workshop on Theoretical Issues in Natural Language Processing, TINLAP 1975, Stroudsburg, PA, USA: Association for Computational Linguistics, 1975, 169 – 174.
- Колковска/Kolkovska 2005:** Колковска, С. Модели на концептуално-семантичните отношения между термините в специален (химически) текст с оглед на автоматичното им разпознаване. // *Литернет*, <https://liternet.bg/publish16/s_kolkovska/modeli/content.htm>. [Kolkovska, S. Modeli na kontseptualno-semantičnite otnosheniya mezhdur terminite v spetsialen (himicheski) tekst s ogleđ na avtomatičното im razpoznavane. // *Liternet*.]
- Ogrodnichuk/Огородничук 2016:** Ogrodniczuk, M., Zawisławska, M. Bridging Relations in Polish: Adaptation of Existing Typologies. // *Proceedings of the Workshop on Coreference Resolution Beyond OntoNotes (CORBON 2016)*. San Diego, USA: Association for Computational Linguistics, 2016. 16 – 22.
- Радева/Radeva 2017:** Радева, В. *Българска лексикология и лексикография*. София: Изток-Запад, 2017. [Radeva, V. *Balgarska leksikologiya i leksikografiya*. Sofia: Iztok-Zapad, 2017.]