

## **ЕЗИК, КУЛТУРА И ДЕЙНОСТ: СЪПОСТАВИТЕЛНО ИЗСЛЕДВАНЕ НА БЪЛГАРСКИЯ И ШВЕДСКИЯ ГОВОРИМ ЕЗИК**

**БИЛЯНА МАРТИНОВСКИ**

ОТДЕЛ ПО КОГНИТИВНА СЕМИОТИКА, ЛУНДСКИ УНИВЕРСИТЕТ  
*bilyanamartinovski@fastmail.com*

Устната реч е фундаментална за човешката когниция и за развитието на човешкия мозък и на обществото, но все още не е достатъчно проучена. Една от причините е, че е трудно да се комбинират качествени и количествени методи за анализ в дългосрочни изследвания по данни за мултимодално общуване. Целта на настоящата статия е да изследва междудейности езикови различия и прилики между говоримия български и шведски език, както и да предложи методи за изследване на връзката между дейност, език и култура в устното общуване. Резултатите от това емпирично проучване показват, че факторът дейност има най-силно въздействие върху устното общуване лице в лице. В статията са представени хибридни качествено-количествени методи и шестнайсет статистически мерки за съпоставителен анализ на говорим език и дейност.

*Ключови думи:* български; шведски; говорим език; култура; дейност; статистика; устна реч; корпусна лингвистика; хибриден метод; когнитивна семиотика

## **LANGUAGE, CULTURE AND ACTIVITY: A CONTRASTIVE STUDY OF COLLOQUIAL BULGARIAN AND SWEDISH**

**BILYANA MARTINOVSKI**

DIVISION FOR COGNITIVE SEMIOTICS, LUND UNIVERSITY  
*bilyanamartinovski@fastmail.com*

Spoken language is fundamental for human cognition and for the development of the human brain and human society yet it is not as rigorously studied as language structure. One of the reasons is the difficulty of combining qualitative and quantitative methods of analysis in longitudinal studies based on authentic multimodal interaction data. The purpose of this paper is to examine cross-activity linguistic differences and similarities of spoken Bulgarian and Swedish and to propose methods for analysis of the relation between activity, culture and language in spoken language interaction. The results of this empirical study show that activity type has strongest impact on face-to-face spoken lan-

guage interaction. The article presents hybrid qualitative-quantitative methods and sixteen measures for cross-linguistic cross-activity spoken language analysis.

*Keywords:* Bulgarian; Swedish; spoken language; culture; activity; statistics; speech; corpus linguistics; hybrid method; cognitive semiotics

## 1. Въведение

Устната реч е от голям интерес не само за лингвистиката, но и за етнографията, историята, когнитивната наука и семиотиката (Pablé & Hutton 2015), за взаимодействието между човек и машина (Kersten 2005), защото има основно значение за развитието на човешката когнитивна дейност (Mac Neilage et al. 1993; Sonesson 2012) и на обществото (Allwood & Hagman 1994). Децата започват да учат и да се развиват въз основа на общуване на говорим език, но това обучение се извършва в рамките на различни дейности. Езикови маркери и черти като комуникативни актове, модалност, интонация, повторения влияят на общуването (Martinovski 2000), но могат да бъдат и продукт на контекстуални фактори (Allwood 1995, Bunt & Black 2000) като дейност (вкл. поддейност, роли, среда), култура и езиков модел. Кой от тези три фактора влияе най-много на комуникацията? Какви методи са най-подходящи за изучаване на влиянието върху езиковата структура, културата и дейността и обратното? Целта на тази статия е да се подходи към тези въпроси чрез прилагане на хибридни качествено-количествени методи и мерки, които могат да помогнат да се разбере как контекстуалните фактори дейност, култура и език влияят върху речевото общуване, което би обогатило разбирането ни за това как говоримият език променя дейностите, културите и езиковата структура във времето. За целта се сравняват български и шведски езикови корпуси от различни видове дейности: както формални дейности в институционална среда, така и неформални видове дейности.

Най-напред се прави описание на теоретичната рамка за комуникативен анализ на дейности, дават се дефиниции на основните понятия и се представят използваните в изследването емпирични данни и методи за анализ. В раздел 5 се представя как два езика реализират една и съща дейност, като се сравняват българският и шведският говорим език в съдебни заседания според маркери като богатство на речника, речево застъпване, обратна връзка, паузи и акцентирание. Раздел 6 съдържа междудейностно междуезиково сравнение на общуване в съдебна зала и беседи на български и общуване в съдебна зала и вечеря на шведски чрез измерване на езикови маркери като обратна връзка, застъпване, паузи, повторения, преформулиране, саморегулация на устната реч, изказвания, думи като знаци и части от речта на всеки език и група дейности. Раздел 7 описва три комплексни мерки за междуезиково и междудейностно изследване на говоримия език, а именно оживеност, стереотипност и предпазливост, последван от заключения в последния раздел.

## 2. Теоретичен модел

Вдъхновен от Витгенщайн (Wittgenstein 1953) и Малиновски (Malinowski 1923), Олууд (Allwood 1976; 1978; 1995) разработва систематичен модел, включващ както структурни, така и контекстуални критерии в една рамка. Дейността се счита за основен организиращ контекстуален аспект на социалната реалност, който влияе върху общуването. Този модел се нарича *дейностен комуникативен анализ* (ДКА) и има следната обща структура:

**Таблица 1.** Дейностен комуникативен анализ (Allwood 1995)

Параметри							
Влияещи				Повлияни			
глобално		локално		глобално		локално	
колективно	индивидуално	колективно	индивидуално	колективно	индивидуално	колективно	индивидуално
цели, роли, страни, артефакти, култура, средства, физически обстоятелства	физически, психически социални характеристики на участниците	поддейности цели, роли, страни	физически, психически, социални характеристики на участниците в поддейност	обща структура на общуване, например, разделения, разговорен ред	средства за общуване, например, движението, жестове, фонология, граматика, речник	поддейности, например, въпрос-отговор, последователност	изказване, принос

Подобно на етнографския комуникативен подход към общуването, моделът на Олууд създава рамка за анализ на различните речеве употреби в различни дейности и в различни общества, като прави таксономично разграничение между неезиковите аспекти на контекста (среда, участници, цели), от една страна, и неговите езикови аспекти, от друга. По този начин той третира като обособени социологическите и езиковите параметри. Динамиката на контекста се разглежда с възможност за промени във влияещите параметри в рамките на една и съща дейност чрез разграничаване между глобални и местни, колективни и индивидуални цели. Моделът е подходящ за анализ на дейности в институционална среда, като например разпити в съдебна зала, където колективните и индивидуалните цели могат да бъдат кооперативни или не и където глобалните и местните аспекти на ситуацията са в стабилни и в същото време ограничаващи отношения.

### 3. Дефиниции

Този раздел предлага дефиниции и описания на основните понятия, използвани в хибридният качествено-количествен анализ на емпиричните данни.

Дейността се определя като специфичен ход на събитията, случващи се в определена обстановка, със специфични правила за поведение, роли и цели.

Културата се определя като динамично свързан набор от исторически, етнически, политически, езикови и географски характеристики, споделени между членовете на дадена група хора.

България и Швеция са приели римското право и инквизиционният принцип в съдопроизводството, но съвременното шведско процесуално право е приело състезателния принцип в съдопроизводството, състезателен формат за съдебни разпити. Освен историческите и политическите фактори, които мотивират решението за приемане на инквизиторска роля на съдията в българските съдебни зали, то е продукт и на недоверие в общността на адвокатите (Терзиев/Terziev 1987). Специфична поддейност, която е от значение за общуването в българския съд, е, че съдията диктува по време на самия разпит или заседание какво да се запише в протокола. В шведския съд се записва всичко, освен ако съдията изрично не се намеси за премахване или коригиране на дадена информация. Интересна особеност е, че и в двете правни култури се счита за положително, ако пледоариите са представени на обективен, ясен език, без широка употреба на реторика (Inger 1986; Malinowski). И в двете страни няма публикации на съдебните речи на видни юристи, каквито има например във Франция и Русия.

Българският език е славянски с типичен нефиксиран словоред, има форми за род и число, темпорална система, включително аорист, и морфологично изразяване на аспекта. На английски българският говорим език се нарича *pro-drop* език, тъй като местоименията могат да се пропуснат, но не *pro-form* език. Така българският е по-синтетичен език в сравнение с шведския. Особено релевантен за съдебния процес е така нареченият доказателствен аспект. Някои форми за минало време носят информация, че споменатото събитие е било възприето или когнитивно преживяно (Fiteva 1998) от говорещия, докато други форми предполагат, че говорещият не е бил свидетел на въпросното събитие.

Шведският език принадлежи към групата на германските езици, без склонение за лице и род. Склонението за мъжки и женски род бе премахнато с решение на парламента, не по естествен път. Шведският език има фиксиран словоред и задължителна позиция на субекта. Аспектът се изразява синтактично, а не морфологично. Разговорният шведски език се характеризира с проформи (изрази в позиция на отговор като „*ja, det gör det*“ – буквално „да, прави/я това“ – като отговор на бинарен въпрос като „Боли ли?“) и различни форми на специален израз за обратна връзка, най-специфичният от които е инхалационният звук, носещ значение на положителна обратна връзка.

Термините *комуникация*, *интерактивност* и *общуване* се използват взаимозаменяемо в тази статия, макар да има различни дефиниции за тях.

Интерактивните повторения и преформулирания са езикови характеристики, които илюстрират хибридният качествено-количествен метод за изследване на реални комбинации от езикови маркери и се дефинират както следва:

*Интерактивно повторение.* Сегмент от речта  $Y$  в изказване  $\beta$  от говорещ  $B$  е интерактивно повторение, ако е напълно идентичен със сегмент от речта  $X$  в изказване  $\alpha$  на говорещ  $A$ , където  $\alpha$  предхожда  $\beta$ , където  $Y$  и  $X$  не се застъпват напълно, където критериите за идентичност се състоят от комбинация от лексика, синтаксис и модалност (но не и интонация).

*Интерактивно преформулиране.* Сегмент от речта  $Y$  в изказване  $\beta$  от говорещ  $B$  е интерактивно преформулиране, ако е частично идентичен със сегмент от речта  $X$  в изказване  $\alpha$  на говорещ  $A$ , където  $\alpha$  предхожда  $\beta$ , където  $Y$  и  $X$  не се застъпват напълно, където критериите за идентичност се състоят от комбинация от лексика, синтаксис и модалност (но не и интонация).

*Интерактивни повтарящи се последователности.* Сегмент от речта  $Y$  в изказване  $\beta$  от говорещ  $B$  и сегмент от речта  $X$  в изказване  $\alpha$  от говорещ  $A$  са интерактивна повтаряща се последователност, ако  $Y$  е интерактивно повторение или преформулиране на  $X$ .

#### 4. Метод и данни

Използваните в настоящото изследване данни за говорим език и дейност се основават на автентични разговори в отделните страни и култури, които съответстват на изследваните езици. Следващият подраздел описва езиковите корпуси и методите за подготовка и анализ на данните.

*Езикови данни и корпуси.* Данните, използвани за това проучване, са част от Гьотеборгския разговорен езиков корпус (ГРЕК), който съдържа 22 вида дейности на шведски език с общо 300 аудио- и/или видеозаписа и 1 221 264 думи. Шведският разговорен корпус за общуване в съдебна зала се състои от транскрибирани и кодирани аудиозаписи на 6 автентични съдебни процеса от общо 34 000 думи. Корпусът за общуване в българската съдебна зала се основава на видеозапис на 5 съдебни процеса в Софийския градски съд през 1995–1996 г. и се състои от 16 000 думи. ГРЕК предоставя транскрипции на автентични записи на други дейности, използвани за междудейностни сравнения в това проучване, като вечери, общуване в работна среда, проповеди, търгове.

Транскрипционният стандарт има за цел да сегментира основните единици за анализ. Процедурата за кодиране идентифицира категориите, които трябва да бъдат анализирани, а в нашия случай предполага прилагането на специално подготвени схеми за кодиране, основани на дефиниция (дефинициите за повторения и преформулирания в раздел 3 по-горе). Стандартът за

транскрипция (Grönqvist & Hashemi-Sofkova 1998), разработен в Катедрата по лингвистика на Гьотеборгския университет, използван в това изследване, включва следните параметри и маркери: раздели (или теми), участници, непълни думи, интонация, удължаване, акцентирание, паузи, застъпване и коментари по конкретни аспекти на речевото събитие. Говорителят се обозначава със знак за долар (\$), двоеточие след инициала на говорещия и интервал. Следните знаци се прилагат към сегментите на думата:

'	глотален стоп с една дума
()	несигурно тълкуване
(xxx)	неразбираема реч или събитие
+	непълна дума (прекъсване, пауза в рамките на думата)
UPPER	подчертан или контрастен акцент
:	удължаване
/	кратка пауза
///	дълга пауза = празнина (gap)
//	неопределена пауза
[N реч] N	припокриващи се квадратни скоби за реч и брой на припокриванията
< >	коментар на интонация, тон на гласа, жестове и т.н.
@	част от коментара.

Българският речник на Николова (Николова/Nikolova 1987) също се използва за сравнение на дейности в това изследване. Корпусът на разговорната реч, който тя използва, се състои от около 100 000 думи, което е около десет пъти повече от българския съдебен корпус, но корпусът на Николова не е транскрибиран и кодиран като ГРЕК, поради което може да се прилага само при определени езикови съставки на дейностите.

*Метод.* Приложени са два различни типа анализ: количествен анализ на съпътстващи прояви на езикови особености и качествен анализ на повторения и преформулировки. Корпусът е количествено анализиран с помощта на програма, разработена в катедрата по лингвистика на Университета в Гьотеборг, наречена ТК-TRASA, която е удобен графичен инструмент за автоматична статистика за транскрипции, предназначен за приложение върху големи корпуси. Той извлича обща статистическа информация въз основа само на транскрипцията, а не на конкретната схема на кодовете.

По-подробен анализ на съвместно възникване на разговорни езикови характеристики, базиран на кодиращи схеми на повтарящи се последователности и комуникативни актове, е проведен чрез формулиране на специфична програма в Prolog, която изпълнява предварително дефинирани търсения и изчисления в кодираните транскрипции с помощта на SICSTUS – интерактивен инструмент за програмиране в Prolog, който е подходящ за подробен анализ на по-малки корпуси.

*Анализ на реални комбинации от езикови маркери.* Анализът на съвместно възникване (Co-occurrence Analysis на английски, Martinovski 2000) се състои от анализ на реални комбинации от езикови черти или маркери, както те реално се проявяват съвместно в речта, и техните функции. Той протича по следния начин: транскрипциите се маркират с набор от маркери и функции като речево застъпване, интонация, комуникативни (речеви и жестови) актове и др., а след това се изброяват и изчисляват съвместите възниквания на различни маркери и функции. Основното предположение е, че различни езикови черти и маркери взаимодействат помежду си при функционирането на различни системи за управление на интеракцията като управление на реда на изказванията, управление на темите, речево саморегулиране, комуникативни актове (речеви актове, които включват невербална комуникация) и др. Следният пример показва извадка от корпуса с транскрибирани и кодирани данни:

### Пример 1.

БТЗ.2: Езикови маркери: речево застъпване ([\_]), повторение/преформулиране (SOURCE\_COPY), потвърждение (CON), разказ (N), декларативно изречение (D), доброволен принос (VOL), WH-въпрос (WHQ), единица за обратна връзка (FB), дума за обратна връзка (FBW), извличане на обратна връзка (FBE), отговор (ANS), оправдание (JUS) и речево саморегулиране (OCM).

1. Свидетел: < [28 **SOURCE22** N D всичко ]28 е прибрано D N **SOURCE22** >  
@ <наречие : обобщително>
2. Съдия: < [28 **COPY22** N FBE D масите ]28 / които (заварихте)  
<прибрахте в склада> D FBE N **COPY22**>  
@ <низходяща интонация>
3. Свидетел: CON FBW да FBW CON OCM ъ.: OCM **SOURCE23** N VOL D  
госпожата си взе: <хладилника> D VOL N **SOURCE23**  
@ <възходяща интонация >
4. Съдия: **COPY23** **SOURCE 24** WHQ <1 <2 кога си взе хладилника  
**SOURCE24** госпожата >2 >1 WHQ **COPY23**  
@ < 1 модалност : въпросителна >1  
@ < 2 нисък тон, низходяща интонация >2
5. Свидетел: ANS D <1 не помня **COPY24** на коя дата си взе хладилника >1  
**COPY24** D ANS JUS D не съм присъствала < 2 когато го е взела >2  
D JUS N VOL D и < 3 билиардната маса >3 също <4 която си я взе  
/// още веднага >4 D VOL N  
@ <1 наклонение : изявително >1  
@ <2 наклонение : преизказно >2  
@ <3 висок тон >3  
@ <4 наклонение : изявително >4

В първото изказване на свидетеля има четири съпътстващи кода на езиковите маркери (с главни букви) – речево застъпване ([ ]), повторение/преформулиране (COPY е името на кода), източник на повторението/преформулирането (SOURCE), декларативно изречение (D) и комуникативен акт разказ (N), плюс коментар, описващ обобщаващата формулировка в изказването. Има първоначално застъпване между изказването на свидетеля на ред 1 и изказването на съдията на ред 2, последвано от кратка пауза от съдията. Застъпване № 28 е едновременна реч, а не повторение, но в същото време е част от частично повторение и преформулиране от съдията на ред 2. Това повторение и преформулиране има две функции: диктовка и формулиране на съдебния протокол и извличане на обратна връзка.

В симултанната реч свидетелят и съдията се позовават на различни елементи в показанията. Свидетелят се позовава на „всичко“, като в същото време съдията се позовава на „масите“ в диктовката си. Изказването на ред 2 може да се нарече и задна обратна връзка (backchannel) в смисъл на частично застъпваща се единица за обратна връзка. На ред 3 свидетелят предлага потвърждение (CON) с дума за обратна връзка (FBW) в начална позиция, с което одобрява формулировката на протокола и определя функцията на изказването на съдията на ред 2 като извличане на обратна връзка. След потвърждението следва звук за речево саморегулиране (OCM – own communication management, Allwood & Hagman 1994), последвано от доброволен принос (VOL), представен като доброволен свидетелски разказ (N) в декларативно изречение (D). Доброволният разказ на свидетеля в декларативно изречение е източник (SOURCE) на повторение/преформулировка от съдията на ред 4. Съдията повтаря думите, но превръща декларативното изречение в полярен въпрос (WHQ). На ред 5 повторението/преформулирането на изказването от съдията на ред 4 се превръща в поредното повторение в отговора на свидетеля на въпроса на съдията на ред 4, последвано от оправдание (JUS). Тази комбинация от съвместно възникващи речеви маркери ANS, D, COPY, JUS, и VOL е позната като защитна реакция или стратегия както в българските, така и в шведските съдебни разпити.

Този пример илюстрира динамичната сложност в реално време на когнитивното и интерактивното формулиране на връзката *функция – маркер* в разговорния език, която изпълнява специфичните цели и ограничения на конкретната дейност.

Когато данните се анализират по подобен начин, може да се наблюдава колко повтарящи се тагове се срещат съвместно с други маркери за функции. По този начин може да се отчете взаимозависимостта на езиковите маркери, техните функции и типичните им комбинации в реално време за всеки говорещ, последователност, дейност, език. Целта е да се получат не само впечатления за това как функционират застъпващите се или повтарящи се последователности и комуникативни актове, но и да се опишат количествено с повече детайли. Този метод на анализ и кодиране е приложен



към българските и шведските съдебни процесуални данни, които са част от ГРЕК.

Тринайсет прости и пет комплексни езикови мерки (таблица 2) се прилагат към езиковите корпуси и се използват при:

- съпоставяне на дейности;
- съпоставяне на езици.

Анализът протича по следния начин:

- Езиково съпоставяне на съдебни заседания на шведски и български език според избрани критерии и мерки.
- Междудейностно междуезиково сравнение на общуване в институционална среда (съдебни заседания) и на двата езика в неинституционални и неформални условия (вечеря на шведски език и среща на чай на български език).

Специфичните критерии за сравнение на говоримия език в списъка по-долу се отчитат автоматично въз основа само на транскрипция: знаци/думи, типове думи/знаци, изказвания, дължина на изказванията, паузи, застъпвания, акцентирание, колокации, части на речта, речево саморегулиране (ОСМ например, звуци на колебание), повторения, преформулировки. Двойките повторения/преформулировки първо се кодират и след това се броят автоматично от програма за анализ на реални комбинации от маркери. Въз основа на тези транскрибирани и кодирани данни се прилагат 16 мерки, описани в таблица 2.

**Таблица 2.** Списък на езикови мерки, приложени върху шведските и българските съдебни заседания

Мярка	Коментар
PauPTok	процент паузи спрямо общото количество думи като знаци в група
UttRC	процент на изказванията в съответната група спрямо общата сума знаци в двете групи
TokRC	процент на знаци в съответната група спрямо общото количество знаци в двете групи
StrPTok	процент на акцентирани думи като знаци по отношение на общото количество думи
OmedPO	процент на застъпване в средна позиция спрямо общия брой застъпвания в езикова група
OiniPO	процент на застъпване в начална позиция спрямо общия брой застъпвания в езикова група
OfinPO	процент на застъпване в крайна позиция спрямо общия брой застъпване в езикова група

OcomPO	процент на пълно застъпване спрямо общия брой застъпвания в езикова група
TovPTok	процент на застъпващи се думи спрямо общия брой думи като знаци в група
OvePUtt	процент на застъпващи се думи като знаци спрямо общия брой думи като знаци в група
Повторения и преформулировки	процент на повторение/преформулиране от общия брой изказвания в група
MLU	Mean Length of Utterance – среден брой думи в изказване
Речник Vocab	“eoretical vocabulary“ measure, Мярка за „Теоретичен речник“
Оживеност	комплексна мярка за оживеност според разговорната езикова дейност
Стереотипност	Комплексна мярка за това колко стереотипна е интерактивната дейност на говорим език. Отразява синтактичната и лексикална еднаквост на изказванията въз основа на видовете колокации от 1–10 думи.
Предпазливост	Комплексна мярка за предпазливост в интерактивната дейност. Стойност на паузи, акцент, саморегулация (ОСМ), обратна връзка, числа, местоимения, застъпване, разделена на общия брой знаци

Българските и шведските съдебни корпуси на говорим език, както са транскрибирани и кодирани, предлагат възможност за много повече мерки за сравнение (Martinovski 2000). Например могат да се измерват езиковите черти и маркери според поддейностите и ролите на участниците или да се изучава реализацията на съмнение и защита в двата езика, както и доминантността на говорещите и равенството според ролите и отношенията за всяка дейност. Мерките в таблица 2 са избрани, за да илюстрират методите за анализ и да дадат по-общ поглед върху връзката между избраните езици и дейности.

## 5. Една дейност на два езика. Съпоставително сравнение на български и шведски говорим език в съдебни заседания

Българският корпус е по-малък от шведския (виж таблица 3 по-долу), затова сравняваме средните стойности, знаейки, че те дават приблизителна картина за изследваните отношения. За да се оцени богатството на речника, може да се измери броят на типовете думи, където наборът от думи се избира произволно (Vocab). Може да се използва средната стойност на брой знаци (думи и други изказки като ‘ъ::’) и процентът на изказванията спрямо целия корпус, както и средният брой думи на изказване, които са свързани с богатството на речника.

**Таблица 3.** Лексикално богатство на българските и шведските езикови данни

Група	Брой знаци	Изказвания	TokRC a.	UttRC b.	MLU c.	Vocab d.	Типове
<b>Български</b>	10533	1115	23.9 %	33.4 %	9.4	2276.6	2345
<b>Шведски</b>	33409	2220	76.1 %	66.6 %	15.1	1920.9	3733

Бележка:

- a. TokRC – процент знаци в съответната група спрямо общото количество знаци в двете групи;
- b. UttRC – процент на изказванията в съответната група спрямо общото количество знаци в двете групи;
- c. MLU – среден брой думи на изказване
- d. Vocab – мярка за лексикалното богатство на речта.

Стойността на MLU по-горе показва, че изказванията на говорещите шведски са по-дълги от тези на говорещите български, въпреки че речникът на българския корпус е по-богат. Това означава, че словоохотливостта на изказванията не допринася за богатството на речника. Обяснението на тази констатация е, че шведският, за разлика от българския, е *pro-form* език, т.е. отговорите и обратната връзка на български, които преобладават в инквизиционните разпити в съдебната зала, трябва да включват конкретния глагол във въпроса, а не проформен израз (вж. раздел 3 по-горе). Както ще видим в следващия раздел 6 обаче, междудейността сравнение между езиките показва, че не езикът, а дейността е определящият фактор за тази комбинация от езикови маркери.

Процентът на паузите и акцентирането на знак в таблицата по-долу показва, че шведските транскрипции от съдебни заседания съдържат повече паузи от българските, но българските съдържат повече акцентирани думи като знаци, въпреки че този по-голям брой паузи може да се обясни отчасти с по-дългите изказвания (стойността на MLU, таблица 3 в шведските съдебни данни).

**Таблица 4.** Пауза и акцент според общата сума словесни знаци в българските и шведските съдебни данни

Група	PaуPTок a.	StrPTок b.
<b>Български</b>	3.9%	6.2%
<b>Шведски</b>	8.4%	4.0%

Бележка:

- a. PauPTok – процент паузи по отношение на общото количество думи като брой знаци в група;
- b. StrPTok – процент на акцентираните думи като брой знаци спрямо общия брой знаци в група.

Освен това процентът на застъпващите се думи като знаци в българските съдебни разпити е четири пъти по-голям, отколкото в шведските. Въпреки че българските изказвания са по-кратки, средно 50,6% от тях се състоят от речеви застъпвания. Обикновено се очаква, че застъпванията в средна позиция са по-чести. Това важи за шведския тип застъпване, което е или медиално, или пълно. Това е илюстрирано в следващата таблица.

**Таблица 5.** Припокриване според позицията в българските и шведските съдебни данни

Мярка Група	OmedPO a.	OiniPO b.	OfinPO c.	OcomPO d.	TovPTok e.	OvePUtt f.
<b>Български</b>	18.7	30.7	30.2	20.1	13.4	50.6
<b>Шведски</b>	26.0	23.6	21.6	28.6	3.1	28.3

Бележка:

- a. OmedPO – процент на застъпвания на средна позиция според общия брой застъпвания в езиковата група;
- b. OiniPO – процент на застъпвания на начална позиция спрямо общия брой застъпвания в езиковата група;
- c. OfinPO – процент на застъпвания в крайна позиция спрямо общия брой застъпвания в езиковата група;
- d. OcomPO – процент на пълни застъпвания спрямо общия брой застъпвания в езиковата група;
- e. TovPTok – процент на застъпващите се думи/знаци спрямо общия брой думи като знаци в езикова група;
- f. OvePUtt – процент на застъпващи се думи/знаци спрямо общия брой изказвания в езиковата група.

Крайното и началното застъпване са по-характерни за българските съдебни разпити, което в дуални разговори означава, че има повече стартове преди предишният говорещ да е завършил изказването си, а също и желание за даване на обратна връзка. Този резултат може да се дължи и на факта, че българските изказвания са средно по-кратки. Сравнявайки двата езика в една дейност според няколко черти на говоримия език като застъпване, паузи, дължина на изказването и лексика, забелязваме, че езиковите различия се определят главно от дейността. В следващия раздел сравняваме двата езика и две различни дейности, за да установим дали дейността има по-силно влияние върху устното общуване от езика и културата.

## 6. Междуезиково междудействие сравнение на устен език

Честотата на частите на речта се очаква да бъде повлияна от спецификата на съответния език. Както е посочено в таблица 6, българските съдебни данни съдържат по-малко местоимения, но това се дължи главно на граматическите особености, а именно, окончанията на съществителните и глаголите включват деиксис за лице (person deixis) и по тази причина намалява броят на местоименията в позиция на подлог. Можем да проверим влиянието на фактора дейност, като използваме речника на Цв. Николова за честотата в неформалната реч (Николова/Nikolova 1987) (последният ред в таблица 6). За целите на това изследване изразите за обратна връзка се считат за специална част от речта.

**Таблица 6.** Междуезиково и междудействие сравнение на средното количество части на речта в съда и на вечеря. Съответните обозначения са **прил** (прилагателно), **нар** (наречие), **мест.** (местоимение), **ОВ** (обратна връзка), **глагол** (глагол), **мем** (междуметие), **съюз, пред** (предлог), **същ** (съществително), **число, РС** (речева саморегулация)

Части на речта	знаци	прил	нар	мест	глагол	ОВ	мем	съюз	пред	същ	число	РС
шведска вечеря	30738	3,8	17,0	23,1	20,8	5,2	0,3	7,3	6,1	11,8	0,5	1,3
шведски съд	33409	3,1	15,3	21,5	19,3	4,8	0,0	8,6	8,6	14,5	1,5	2,07
български съд	10533	2,2	4,7	11,9	30,1	2,8	0,0	10,0	10,1	23,0	3,3	1,7
български беседи	100000	3,8	8,8	18,3	22,8	7,5	1,3	10,4	8,0	15,8	3,2	1,3

Наблюденията показват, че дейността понякога е по-силен фактор, а понякога по-силно влияние има езикът. Прилагателните, наречията, местоименията, обратната връзка и междуметията намаляват в съдебните данни и в двата езика, а останалите части на речта се увеличават. Изключение правят глаголите и съюзите в българския език. Даването и изискването на обратна връзка е по-често срещано в българската беседа в сравнение с шведската вечеря, но намалява повече в българската съдебна среда в сравнение с шведската съдебна беседа. Това е свързано с диктовката и с инквизиционния принцип на съдопроизводството, според който съдията контролира реда на изказванията. Изглежда обаче, че този резултат е повлиян и от пуристското отношение, че думите за обратна връзка, като „бе“, „де“, „а“ (използвани само като подбуда на обратна връзка, подобно на шведските „va“, „какво каза“ и „не мислиш ли“), които са най-често срещаните думи в ежед-

невната устна реч (Николова/Nikolova 1987: 24), са паразитни изрази и част от така наречения „уличен език“, т.е. обозначават по-нисък регистър, културно или прагматично непредпочитан в институционална среда.

Материалът от български беседи, като среща на чай например, показва, че високата честота на застъпванията е специфична характеристика на съдебната обстановка, а не на българската интерактивната култура като цяло, тъй като процентът на застръпени знаци в беседата е дори по-нисък от застъпванията в шведските съдилища.

**Таблица 7.** Обратна връзка, повторение/преформулиране и застъпвания в българските беседи и съдебни разпити и в шведските вечери и съдебни разпити

Мярка Дейност	Изказвания	Брой знаци	Обратна връзка а.	Повторения/ преформ. % от изказания б.	ТовРТок % с.
български съд	1115	10533	4.8	16.4	13.4
български беседа	470	5966	6.3	9.3	2.5
шведски съд	2220	33409	4.8	6.5	3.1
шведска вечеря	3251	30738	5.2	–	6.5

Бележка:

- а. Процент на обратна връзка изразни знаци спрямо общата сума знаци;
- б. Процент на повторения/преформулиране спрямо общия брой изказвания;
- с. Процент на застъпени думи като брой знаци спрямо общата сума думи като брой знаци в група.

Горепосочените данни предполагат, че интерактивното повторение или преформулиране е по-предпочитана форма на обратна връзка в съдилищата в сравнение с по-неформалните разговори като беседите. Докато повечето от неповтарящите се изрази за обратна връзка по време на беседите са положителни, повече от половината от повторенията/преформулираната са част от разногласия или отрицателна обратна връзка. Можем да заключим, че българските беседи показват по-силна нужда от положителна обратна връзка и повратна координация в сравнение със съдебните заседания. Нямаме данни за честотата на повторенията/преформулировките в шведската вечеря и затова не можем да заключим по-категорично, че интерактивното повторение/преформулиране е по-често срещано в българския език, отколкото в шведския разговор, както се предполага в петата колона по-горе. Подобно на българските данни шведската вечеря съдържа повече обратни

връзки (с изключение на повторенията/преформулиранията) в сравнение с шведските или българските разпити в съдебна зала. Обратното важи за застъпванията: говорещите се застъпват повече в българските съдилища отколкото в приятелски разговори, които са неформални интеракции, докато говорещите се застъпват по-малко в шведските съдилища, отколкото във вечерите. Методите, използвани тук, ни позволяват да наблюдаваме точни отношения, ако данните са напълно сравними, и да наблюдаваме тенденции, ако данните не са напълно сравними, както е в този случай.

Средната честота на частите на речта, включително обратната връзка, речевата саморегулация и процентът на застъпване и повторение/преформулиране показват, че дейността има голямо влияние върху усното общуването както на български, така и на шведски език. Нека видим какво могат да покажат комплексните мерки за междуезикова междудействителна интерактивност за връзката между езика и дейността в общуването.

## 7. Комплексни мерки за междуезиково и междудействително описание на разговорния език

За да се проведат междудействителни и междуезикови сравнения в големи разговорни корпуси, могат да се разработят по-комплексни мерки (критерии), които след това да се приложат към целия езиков корпус. Allwood & Hagman (1994) разработват мярката за оживеност, за да „уловят различията между социалните дейности в разговорния език“ (Allwood & Hagman 1994: 14). Те изискват достъп до различни видове езикови маркери, които се основават само на транскрипции. За да се съпоставят две дейности (съдебни разпити и беседи/вечери) на два езика – български и шведски – се прилага мярката оживеност и се разработват още две, а именно мярка за стереотипност и мярка за предпазливост в общуването, които са подходящи за съпоставяне със съдебни данни.

*Интерактивна оживеност.* Тази мярка се състои от прости мерки като акцент (StressPTok – процентът на акцентирани думи като брой знаци спрямо общото количество думи като знаци), речево застъпване (ToOVPtok – процент на застъпващи се думи като брой знаци спрямо общото количество думи като знаци), OvePUtt – процент на застъпващите се думи като брой знаци спрямо общия брой изказвания в група), средна дължина на изказване (MLU – средната стойност на броя знаци в изказване) и паузи (PausPTok – процентът на паузите като брой знаци по отношение на общото количество думи като знаци). Приложената формула е:

$$\underline{(2. ToOVPtok) + OvePUtt + (2.StressPTok) - (2.PausPTok) - MLU}$$

**Total number of tokens.**

Според тази формула колкото повече са застъпванията и акцентиранията и колкото по-малко и по-кратки са паузите и изказванията в дадена дейност, толкова по-оживена е интерактивната езикова дейност. Обратно, колкото по-чести и по-дълги са паузите и изказванията и по-малко са застъпванията и акцентите, толкова неоживена е дейността.

*Стереотипност.* Тази мярка отразява синтактичната и лексикална еднаквост на изказванията въз основа на видовете колокации. Тя отразява повторяемостта на комбинации от думи, започващи от една дума и завършващи с десет думи в колокацията. Всички дейности са склонни да бъдат по-стереотипни при по-кратки колокации и стават по-малко стереотипни с увеличаването на броя на думите в колокацията. Критична точка е колокацията, състояща се от пет думи, защото в този момент най-драматичните промени в мярката „се случват“ в повечето от 22-те дейности в корпуса. Тези, които нямат нулева стойност, например за десет нива на колокацията, са най-стереотипните. В последната колона давам само стойностите, които съответстват на повторяемостта на една дума (пред наклонената черта) и на колокациите от десет думи (след наклонената черта).

**Таблица 8.** Сравнение между езици и дейности според оживеност и стереотипност

Комплексни мерки Група	Оживеност	Стереотипност
Български съд	71.2	37.7/2.1
Шведски търг	-70.3	35.7/0.02
Шведски съд	9.4	25.4/0.0
Шведска вечеря	91.2	26.6/0.0

Според данните българските съдилища са далеч по-стереотипни от шведската вечеря или съдебните заседания и в това отношение са подобни на търговете, както е показано на фигура 1 по-долу. Всъщност за това допринасят инквизиционната рутина и повторенията/преформулировките в диктовката, чиято функция може да се разглежда и като мярка срещу недоумения.

Според мярката за оживеност българското съдебно заседание прилича повече на шведска вечеря, отколкото на шведско съдебно заседание, главно защото съдържа повече видове думи, застъпващи се и акцентирани думи и по-малко паузи.





**Фигура 1.** Междудейностна и междузикова скала, основана на мерки за жизненост и стереотипност

За да проучим по-нататък въздействието на дейността, можем да сравним мярката за оживеност и стереотипност според видовете дейности. Шведските данни обхващат по-широк спектър от дейности за сравнение от българските данни, поради което можем да сравним много дейности само на шведски език. Таблица 9 показва резултатите за шест от двадесет и двете дейности на шведски в корпуса ГРЕК. Последната колона дава само стойностите, които съответстват на повторемостта на една дума (пред наклонената черта) и на колокациите от десет думи (след наклонената черта):

**Таблица 9.** Междудейностно сравнение според мерките за стереотипност и оживеност в шведските данни

Комплексна мярка Дейност	Среден брой говорещи	Оживеност	Стереотипност
Всички	9	30.1	25.178 / 0.0
Проповед	5	-82.6	18.5/0.3
Търг	7	-70.3	35.7/0.02
Съд	6	9.4	25.4/0.0
Съвместна работа	3	11.3	28.15/0.025
Формална среща	11	33.9	27.7/0.0
Формална вечеря	9	91.2	26.6/0.0

Проповедта е една от най-малко оживените дейности, защото е най-монологична. Търгът, който се състои основно от оферти и потвърждения, също не е много оживен. Дейността, ориентирана към решаване на работни задачи, не е много оживена, но е по-стереотипна от проповедта. Всъщност търгът изглежда е най-стереотипната дейност, с разпределение на единни колокации по отношение на колокации от пет думи (но дори и на колокациите от десет думи). Съдебните заседания, включително разпити, четене на обвиненията и окончателните пледоарии, не са нито изключително стереотипни, нито много оживени. Формалната среща и вечерята са по-стереотипни от съдебните разпити и по-малко стереотипни от търга, но разговорите на вечеря са далеч по-оживени от съдебните разпити и формалната среща. Тя е подобна на разговорите на вечеря по отношение на стереотипното използване на синтактични конструкции, но е много по-оживена.

Можем да кажем, че ако си представим дейностите като континуум по скалата на оживеност и стереотипност, търгът е на единия полюс, а вечерята – на другия. Разговорът в съда е по средата на тази скала с лека тенденция към търга. Тъй като разговорът на вечеря се очаква да бъде различен от институционални ситуации (Atkinson&Drew 1979; Ilie 1999) и тъй като това е най-оживената дейност в корпуса, можем да го сравним с разпитите. По този начин можем да получим някои общи характеристики на устната реч в съда. Гореспоменатите по-комплексни мерки показват, че разговорът на вечеря съдържа повече от осем пъти повече застъпени знаци от съдебните данни, но има много по-малко паузи и акцентирани думи.

Решенията, взети от съда, могат да бъдат фатални, което предполага, че участниците са склонни да бъдат вербално по-предпазливи в съдебните зали, отколкото на вечери. Следователно предпазливостта би била полезна за сравнение и между езиците. Очакваме тя да има сходни стойности на български и шведски поради влиянието на сериозността на дейността. Следващият подраздел формулира такава мярка.

*Интерактивна предпазливост.* Може да се предположи, че по-големият брой колебания означава по-голяма грижа при формулирането на изказванията (Martinovski 2000). Същото може да се каже и за по-големия брой паузи в съда. И двете характеристики насочват към по-внимателен стил на общуване. Тъй като предизвикването на обратна връзка е едно от основните разговорни езикови средства за съвместна реализация на комуникативни проекти, в които има нужда от постоянна проверка на това, което се възприема и/или разбира, както и от търсенето на внимание (особено по въпроси от голямо значение като показанията), по-голямото количество изрази за обратна връзка може да бъде индикация за по-силна нужда от такава координация, а следователно и за интерактивна предпазливост. Използването на числа, суми и подобни, както и на лични имена вместо местоимения също може да бъде индикация за предпазливост. Така можем да създадем комплексна мярка за това колко предпазлив е разговорният език. Колкото повече

са речевите саморегулации (ОСМ в пример 1 по-горе), паузите, акцентиранияте думи, изразите на точност и изразите за обратна връзка, толкова по-предпазлива и прецизна е устната интеракция. Колкото повече са застъпванията и местоименията (вместо лични имена), толкова по-непредпазлив е устният език. Така формулата на комплексната мярка за интерактивна предпазливост в дадена дейност е следната:

$$\frac{(2. \text{ Паузи}) + \text{Акценти} + (2. \text{ ОСМ}) + (2. \text{ Обратни връзки}) + \text{Числа} - (2. \text{ Междуимения}) - (2. \text{ Застъпване})}{\text{обща сума на знаците}}$$

Приложена върху данните, тази формула показва, че говоримият български е полярно краен и в двете посоки в двата различни типа дейност в сравнение с шведската устна реч.

**Таблица 10.** Сравнение между езици и дейности според предпазливост

Език / Дейност	Интерактивна предпазливост
български съд	- 0.02
шведски съд	- 0.1
шведска вечеря	- 0.7
българска беседа	1,1

Съдебните разпити изглеждат по-предпазливи от разговорите на вечеря независимо от езика. Те могат да се тълкуват като по-прецизно ориентиран тип вербална дейност. Предпазливостта е свързана със задължителност, която дейността предполага. По този начин предпазливостта и прецизността в позоваванията (големия брой числа и лични имена) е отражение на целта на тази дейност, а именно обективно вземане на решения, основано на справедливост и честност. Ето защо не е изненадващо, че въпреки факта, че мярката за предпазливост не отчита количеството повторения/преформулировки, нито видовете речеви актове, българското съдебно заседание е описано като по-предпазливо от шведската вечеря.

Формулата за предпазливост използва обратна връзка и местоимения, които са сравнително слабо представени в българските съдебни данни (вж. също таблица 6 и таблица 7) и влияят върху крайната мярка. Използването на обратна връзка допринася за предпазливостта, а използването на местоимения показва по-малко предпазливост. Местоименията се използват по-малко, тъй като българският изразява морфологично лице и е pro-drop език. Обратната връзка се използва сравнително по-малко в българския говорим

език независимо от дейността. Тя допринася за по-ниския резултат в мярката от български съдебни заседания в таблица 10, което предполага, че и езиковият фактор влияе върху тази мярка. Тъй като неформалната българската дейност беседа, има много по-ниска стойност за предпазливост от българските и шведските съдебни заседания, дори по-ниска от шведската стойност за езикова и интерактивна предпазливост на вечеря, може да се заключи, че факторът дейност наистина влияе най-силно за този набор от данни независимо от езика.



Фигура 2. Междудейностна и междуезикова скала според мерки за стереотипност, предпазливост и оживеност

## 8. Заключение

В статията се разглежда устното езиково общуване чрез разработване на мерки, които имат за цел да допринесат за развитието на българската и шведската прагматика, семиотика и лингвистика и да обогатят разбирането за говоримия език и връзката му с контекстуални фактори като култура, дейност и структура на езика. Прилагането на мерките върху междуезиковите и междудейностни данни показва, че автоматизираният лингвистичен анализ, основан на хибридна качествено-количествена методология и анализ на съвместно възникване на реални комбинации от езикови маркери, е полезен инструмент за съпоставително описание на вида дейност и говоримия език.

Проучването установява, че факторът дейност влияе върху взаимодействието лице в лице на говорим език повече от езиковия фактор или фактора национална култура, което подкрепя прогнозите на теоретичния модел на

Олууд за дейностен комуникативен анализ. Ако дейността е по-въздействаща от културата и езика, тогава културните и езиковите промени се влияят най-вече от дейностите, в които хората участват, много от които се основават на говоримия език.

Изследването показва, че хибридната качествена и количествена методология е от голямо значение. За да се изследва връзката между език, култура и дейност, е необходимо да се сравнят стойностите на данните както междудейностно, така и междуезиково. Например лингвистичното обяснение на по-честата употреба на местоимения в шведския (защото не е pro-drop език) и по-кратките изказвания с по-малко думи в българския говорим език (защото е pro-drop и по-синтетичен език) е опровергано в междудейностния анализ. Там виждаме, че стойностите на тези езикови маркери на двата езика се променят в унисон с дейността, а не според типа език или културата. Друг пример е застъпването. Данните от българските беседи показват, че високата честота на застъпване в съдебна обстановка се дължи на формата на дейността, а не на интерактивната култура като цяло, тъй като средният брой на застъпени знаци в беседата е по-нисък дори от застъпването в шведските съдилища. Българският съдебен процес е инквизиционен, шведският е състезателен, но основната цел на съдебното процедура е отразена в комбинацията от разговорни езикови маркери, като се изтъква подчертаният стремеж към предпазливост и обективност на дейността независимо от културата и езика.

Ограниченията на изследването са в малкия обем на някои от данните, голямата разлика между обема на някои данни и това, че не всички данни са транскрибирани и кодирани с едни и същи стандарти и схеми. Бъдещи семиотични и лингвистични изследвания на устната речева интеракция на различни езици в различни култури и дейности могат да развият по-богати и актуализирани корпуси на българския говорим език, за да предоставят по-прецизни и подробни познания както за него, така и за други говорими езици, като използват и доразвиват данните, детайлната статистическа методология и езиковите мерки, представени в тази статия.

**Благодарности:** Благодаря на проф. Максим Стаменов за съдействието му при превода на статията на български.

#### ИЗТОЧНИЦИ / SOURCES

Николова 1987: *Николова, Цв.* Честотен речник на българската разговорна реч. София, Наука и изкуство. [Nikolova 1987: *Nikolova, Tsv.* Chestoten rechnik na balgarskata razgovorna rech. Sofia, Nauka i izkustvo.]

#### ЛИТЕРАТУРА / REFERENCES

Allwood 1976: *Allwood, J.* Linguistic Communication as Action and Cooperation. Gothenburg Monographs in Linguistics 2. University of Göteborg, Department of Linguistics: Gothenburg.

- Allwood & Hagman 1994: *Allwood, J., J. Hagman*. Some Simple Automatic Measures of Spoken Interaction. – In: *Proceedings of the XIV Conference of Scandinavian Linguistics & Gothenburg Papers in Theoretical Linguistics 72*, University of Gothenburg: Gothenburg.
- Allwood 1995: *Allwood, J.* Activity based approach to Pragmatics. Gothenburg Papers in Theoretical Linguistics 76, University of Göteborg, Department of Linguistics: Gothenburg.
- Atkinson & Drew: 1979 *Atkinson, J. M., P. Drew*. Order In Court. The Organisation of Verbal Interaction In Judicial Settings. MSSRC: London.
- Bunt & Black 2000: *Bunt H., W. Black*. Abduction, Belief and Context in Dialogue. John Benjamins Publishing: Amsterdam-Philadelphia.
- Fiteva 1998: *Fiteva, M.* Evidentiality in Bulgarian. 6<sup>th</sup> International Pragmatics Conference, Reims, France.
- Grönqvist & Sofkova-Hashemi 1998: *Grönqvist, L., S. Sofkova-Hashemi*. Towards Multimodal Spoken Language Corpora. – In: *Proceedings of the Workshop on Partially Automated Techniques for Transcribing Naturally Occurring Continuous Speech – Transcribe '98*, Quebec, Canada.
- Ilie 1999: *Ilie, C.* Question – response argumentation in talk shows. *Journal of Pragmatics*. 3(8), pp. 975–999.
- Inger 1986: *Inger, G.* Svensk Rättshistoria. Liber Förlag: Lund.
- Kersten 2005: *Kersten, G.* E-negotiation systems: Interaction of people and technologies to resolve conflicts. *The Magnus Journal of Management* 1(3) pp. 71–96.
- MacNeilage et al. 1993: *MacNeilage, P., M. Studdert-Kennedy, B. Linblom*. Hand Signals. – In: *The Sciences. The Academy of Sciences*, New York.
- Malinowski 1923: *Malinowski, B.* The problem of meaning in primitive languages. – In: Ogden C.K. & LA Richards L.A. (Eds.) *The Meaning of Meaning*. Harcourt, Brace & World, New York, pp. 296–336.
- Martinovski 2000: *Martinovski, B.* The Role of Repetitions and Reformulations in Court Proceedings – a Comparison of Sweden and Bulgarian. Gothenburg Monographs in Linguistics. University of Gothenburg: Gothenburg.
- Pablé & Hutton 2015: *Pablé, A., C. Hutton*. Signs, Meaning and Experience: Integrational Approaches to Linguistics and Semiotics. De Gruyter Mouton: Boston/Berlin.
- Sonesson 2012: *Sonesson, G.* The foundation of cognitive semiotics in the phenomenology of signs and meaning. *Intellectica* 58(2), pp. 207–239.
- Terziev 1987: *Terziev, V.* Sadebnata zashita i advokaturata. Sofia, Nauka i izkustvo. [Terziev 1987: *Terziev, B.* Съдебната защита и адвокатурата. София, Наука и изкуство.]
- Wittgenstein 1953: *Wittgenstein, L.* Philosophical Investigations. Blackwell: Oxford.

✉ Доц. д-р Биляна Мартиновски / Assoc. Prof. Bilyana Martinovski, PhD  
Division for Cognitive Semiotics, Lund University  
Kungshuset (The King's House) Kyrkogatan 8  
Lund, Sweden