

АРГУМЕНТИ ЗА ДОБРО ИЛИ ЛОШО: УСЕТ КЪМ АРГУМЕНТНАТА СТРУКТУРА И ПРОЦЕСЪТ НА РАЗБИРАНЕ НА ИДИОМАТИЧНИ ИЗРАЗИ ПРИ РАЗЛИЧНИ ПОПУЛАЦИИ

МИЛА ВЪЛЧАНОВА, ВАЛЕНТИН ВЪЛЧАНОВ

ЛАБОРАТОРИЯ ПО УСВОЯВАНЕ И ОБРАБОТКА НА ЕЗИКА

НОРВЕЖКИ НАУЧНО-ТЕХНОЛОГИЧЕН УНИВЕРСИТЕТ – НОРВЕГИЯ

mila.vulchanova@ntnu.no; valentin.vulchanov@ntnu.no

В статията се разглеждат подходите към аргументната структура на глаголите от гледна точка на информацията, която се кодира лексикално в значението им. Авторите изследват начините, по които усетът на носителите на даден език към аргументната структура може да се тества експериментално при различни типове изрази, включително идиоми, и представят изводите от актуални емпирични изследвания в тази област.

Ключови думи: лексикално кодирана информация, глаголи, аргументна структура, езиков усет, разбиране на идиоматичните изрази

ARGUMENTS FOR GOOD OR BAD: SENSITIVITY TO ARGUMENT STRUCTURE AND IDIOM PROCESSING ACROSS POPULATIONS

MILA VULCHANOVA, VALENTIN VULCHANOV

LANGUAGE ACQUISITION AND LANGUAGE PROCESSING LAB

NTNU – NORWAY

mila.vulchanova@ntnu.no; valentin.vulchanov@ntnu.no

This paper addresses approaches to verb argument structure from the point of view of the information which can be assumed to be lexically encoded in the verb. It explores ways in which speakers' sensitivity to verbs can be investigated experimentally across types of expressions, including idioms/non-literal language, and reports findings from recent empirical research in that domain.

Keywords: lexically encoded information, verbs, argument structure, sensitivity, idiom processing

1. Защо точно аргументната структура?

Аргументната структура на глаголите заема централно място в съществуващите теоретични подходи, изучаващи природата на лексикалния за-

пас и взаимодействието му със синтактичната структура. В зависимост от лингвистичната традиция и изследователската парадигма в литературата са се наложили редица термини за означаване на различни варианти на това явление: субкатегоризация, тета-решетка, глаголна валентност, лексикална спецификация. Фокус на научните дискусии и разногласия обикновено е начинът на представянето на информацията в речника и въпросът дали и в каква степен лексикалната информация си взаимодейства с механизмите на граматиката, или, напротив, се приема, че тези два компонента не са свързани и функционират независимо един от друг. По същество така поставеният изследователски проблем се заключава в това дали като част от лексикално кодираната информация аргументите на даден глагол предопределят структурата на изреченията, чието предикативно ядро е той. Във връзка с това Джакендоф предлага теоретичен модел, състоящ се от три относително автономни нива на описание на езиковата структура (фонологично, синтактично и семантично), между които съществуват отношения на активно взаимодействие, регулирани от правила за съотнасяне (Jackendoff 2002). Преобръщането на формулировката на въпроса поставя фокуса върху това кои езикови аспекти е нужно да се съхраняват в дългосрочната памет, и кои – не. Тази смяна на изследователския ракурс полага основите на моделите на езика в сферата на изучаването и обработката му и осветлява езиковата структура от гледна точка на ролята на основните когнитивни механизми, подпомагащи езиковите процеси, в усвояването и разбирането на езика и създаването на реч (Pinker, Jackendoff 2005). Джакендоф предлага радикално разбиране за думата като единица, съвместяваща трите вида структура (фонологична, граматична и семантична), която се представя под формата на многократни правила за съотнасяне, осъществяващи връзката между фонологичната форма, граматичната (морфологичната структура и лексикалната категория) и семантичната информация. „Следователно думата представлява съхранявана от мозъка асоциация между три вида структури, съотнесени помежду си с помощта на еднакъв индекс. Иначе казано, речникът се разглежда не като отделен езиков модул, който служи като вход за синтактичния компонент, а като част от компонентите на взаимодействието между отделните структури: той способства осъществяването на връзката между фонологичната форма, синтаксиса и семантиката“ (Jackendoff 2002: 27–28).

Други модели в рамките на същата езиковедска традиция обикновено вземат предвид главните особености на думите, които имат отношение както към усвояването и съхраняването им в дългосрочната памет, така и към активната им употреба (например т.нар. лексикален достъп). Обща черта на тези постановки е, че информацията за характера и броя на аргументите представлява част от семантичната репрезентация на глаголите. В широко утвърдия се модел на Левелт (Levelt 1993) аргументите се възприемат като част от значението на лемата (основната форма) и се представят като аргументи на концептуални функции. Близката по дух, макар и

самостоятелно формулирана хипотеза за лексикалното качество (Lexical Quality Hypothesis) разглежда лексикалните единици като единство на фонологична, правописна и семантична информация (Perfetti, Hart 2002; Perfetti 2007). Изграждането на тази обща ядрена (лексикална) репрезентация на думата се основава на многократното ѝ срещане в процеса на усвояване на лексиката. Бързото и ефективно извличане на дадена дума от паметта се основава на качеството на съвкупността от характеристики, които формират тази репрезентация. При чуването на определена лексема (и възприемането на нейния акустичен образ) се извършва едновременното активиране на отделните компоненти на значението ѝ, включително на знанието за броя и типа на аргументите, когато думата е глагол (Koenig et al. 2003). Важен аспект на модела, предложен от Кьониг и колектив, е, че лексикално кодираната информация няма пряко отношение към въпроса дали аргументите трябва да получат синтактична реализация, или не. По същество това предполага, че въпреки че представянето на определени аспекти на информацията за аргументите има задължителен характер, те могат да останат и фонологично неизразени. Въз основа на тази информация се предопределя преценката на носителите на даден език относно приемливостта на граматичните структури, а в процеса на езиковата обработка – и способността им да предскажат езиковия материал, който може да се появи в определен контекст. Може да се заключи, че по този начин лексикалната информация, от една страна, подпомага езиковата обработка, като дефинира свободните (тематичните) позиции, които трябва да бъдат запълнени, а от друга – задава ограничения върху езиковата структура/граматичния компонент чрез налагане на определени изисквания по отношение на броя и семантичните характеристики на единиците, които може да се реализират в тези позиции (вж. подробната дискусия и предложения теоретичен модел у Димитрова-Вълчанова (Dimitrova-Vulchanova 1999), както и изложението у Димитрова-Вълчанова и Вайсгербер (Dimitrova-Vulchanova, Weisgerber 2007).

2. Какво показва чувствителността към аргументната структура?

С оглед на централната роля, която заема репрезентацията на думите в съзнанието на носителите на езика, и факта, че представлява част от лексикалната информация, активирана при срещането на дадена лексема в потока на речта, може да се приеме, че аргументната структура, т.е. знанието за броя и характера на глаголните аргументи, е отражение на езиковата компетентност и следователно може да се използва в експериментални изследвания като косвен индикатор за езиковите умения. Тъй като бързото и ефективно извличане на дадена дума се основава на качеството на отделните характеристики, които съставляват част от репрезентацията ѝ, чувствителността към аргументната структура на глаголите при лицата, за които даден език е роден, е показател за качеството на съхраняваната лексикална информация и за интеграцията на дадения глагол в лексикалната мрежа, изграждаща менталния лексикон.

Като се изхожда от предположението, че усвояването на думите се основава на сходни когнитивни механизми и води до аналогични резултати и при изучаващите даден език като чужд, чувствителността към аргументната структура на глаголите може да се използва като тест за проверка на лексикалната консолидация и интеграция и при лицата от тази популация. Носителите на даден език и усвояващите го като втори се различават в едно важно отношение, а именно по това, че формираното лексикално знание по родния език често възпрепятства усвояването на нови думи от чуждия. Според някои изследователи при представителите на втората група се наблюдава т.нар. пълен пренос на лексикално съдържание от майчиния език към еквивалентните думи в чуждия – явление, залегнало в основата на хипотезата за пълния пренос (*Full Transfer Hypothesis*, Stringer 2010). С течение на времето при засилено съприкосновение с втория език усвояващите го започват да приспособяват лексикалната репрезентация на думите, която съхраняват в съзнанието си, и да „реконцептуализират“ ситуациите, означавани от глаголите в чуждия език (Treffers-Daller, Calude 2015). Съществуващите изследвания сочат, че чувствителността към аргументната структура може да се използва като индикация за равнището на владееене на втория език и че в това отношение се регистрират значителни разлики между изучаващите го в начална и в по-напреднала степен (Larrañaga et al. 2012; Treffers-Daller, Calude 2015; Treffers-Daller, Tidball 2015; Johnsen 2016; Reine 2016). Цитираните разработки демонстрират, че често в ранните етапи на усвояване се наблюдава пренос на семантични особености, характерни за глаголите от родния език в езика цел, включително на специфики, свързани с кодирането на аргументите и лексикализацията на участниците в ситуациите, описвани от глаголите. Споменатите научни изследвания демонстрират още и че по-високото ниво на владееене на даден втори език и честото съприкосновение с него може да подпомогнат избягването на „капана“ на езиковия пренос (Cabrera, Zubizarreta 2005; Stringer 2010).

3. Изследване на чувствителността в родния и в чуждия език

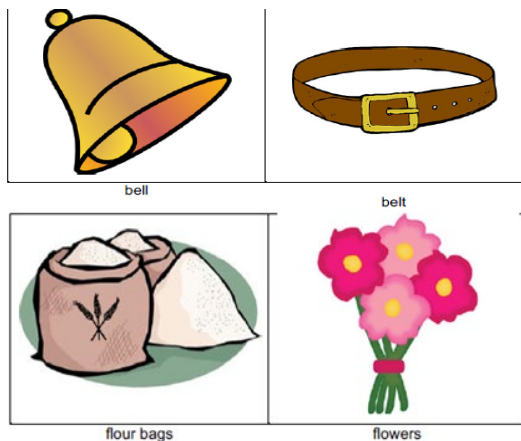
Съществуват различни методи за експериментално изследване на чувствителността към аргументната структура. В разработките върху усвояването на втори език, споменати по-горе, на участниците са поставяни задачи за създаване на свободен описателен текст, а получените отговори са кодирани в подходящ за по-нататъшния анализ формат. В настоящата статия обект на интерес е въпросът в каква степен тестваните глаголи са употребени в синтактичен контекст, характерен за езика цел. Резултатите са сравнени с отговорите, получени от носители на езика при изпълнение на същата задача. Като алтернативна стратегия при този тип експерименти се използва проследяване на поведението на участниците при слухово възприемане на изречения стимули на езика цел, съдържащи изследваните

глаголи. Този тип експериментален дизайн позволява оценяването на степента, в която изучаващите даден втори език очакват и са в състояние успешно да предвидят именната фраза, която ще запълни дадена аргументна позиция в тестовото изречение при възприемането му в реално време. Именно такъв експериментален модел, имплементиран в съответствие с т.нар. парадигма на зрителния свят (*Visual World Paradigm, VWP*) (Allorenna et al. 1998; Altmann 1997; Altmann, Kamide 1999), е използван в описаните по-долу експерименти с цел да се провери хипотезата установяват ли се различия в способността за предвиждане на глаголните аргументи при групи носители на норвежки, владеещи английски на различно (пониско и по-високо) ниво.

Метод на изследване

В изследването участват 25 16-годишни ученици от гимназиалния курс (групата на владеещите английски на средно ниво) и 26 студенти (представители на възрастовата група 20 – 28 години, със средна възраст 24,07 г.). Задачата се състои в слухово възприемане на изречения на английски език, по време на което се извършва проследяване на погледа на участниците по посока на зрителните стимули, показвани на компютърните екрани пред тях. Като стимули са използвани: изображение на именната фраза, която запълва аргументната позиция в изречението, фонологична кохорта (дума, започваща със същата последователност от фонемни като целевата фраза) и две изображения дистрактори (вж. фигура 1). Данните са записани с помощта на устройство за проследяване на погледа модел Tobii T120 при честота на дискретизация 60 Hz. В подготвителната фаза са събрани данни от 105 носители на английски език (представители на възрастовата група 15 – 45 години със средна възраст 19,75 години) с помощта на тест за попълване на пропуснати думи (*cloze test*), съдържащ 77 често срещани английски глагола, които изучаващите английски като чужд би трябвало да владеят. Участниците с роден език английски изпълниха тестовите задачи с помощта на онлайн формуляр. Тестовите за попълване на пропуснати думи се използват за оценка на вероятността определени лексеми да се появят след целевата дума. Например в изречението *John kicked the ----* („Джон ритна ----“) обект на интерес представлява коя именна фраза е най-вероятно да се появи непосредствено след глагола *kick* („ритам“). Въз основа на процентното съотношение на англоезичните участници, които посочат конкретен отговор, може да се изчисли вероятността за появата на всяка от възможните запълващи фрази в съответната позиция (*cloze probability*). Както според данните, получени от тестваните носители на английски, така и според направената корпусна извадка като запълващ елемент на позицията вдясно от глагола *kick* най-често се появява съществителното *ball* („топка“). Това предполага, че когато чуят тестово изречение, чието предикативно ядро е глаголът *kick*, участниците, за които английският език е роден, ще насочат погледа си към изображението на топка по-

рано, отколкото към някои от останалите обекти, изобразени на екрана. Тази реакция е показателна, че тестваните лица очакват в позицията на пряко допълнение да се появи именната фраза *the ball* („топката“).



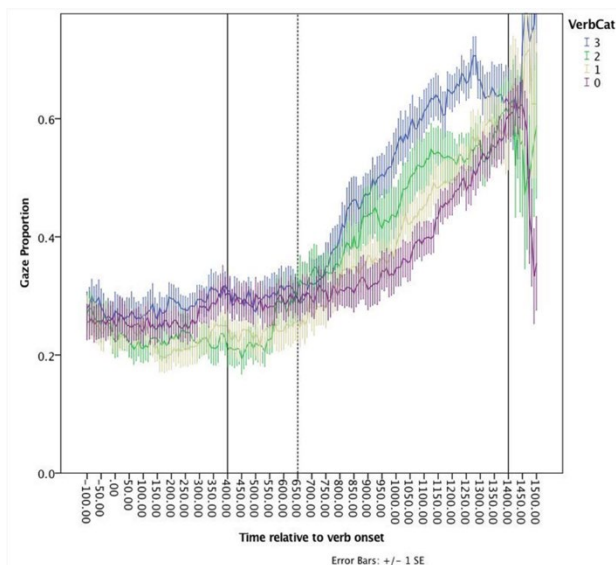
Фигура 1. Пример за представяне на избраните стимули за изречението *Jesse fastened the belt eagerly* („Джеси затегна колана нетърпеливо.“), които включват изображение, съответстващо на целевата дума (*belt* – „колан“), неин фонологично близък конкурент (*bell* – „звънец“) и два дистрактора (*flowers* – „цветя“ и *flour bags* – „чували с брашно“).

За тестването на носителите на английски език са използвани множество от предварително подбрани глаголи. Трийсет и пет от тях са определени като ограничени по отношение на лексикалната селекция на именните фрази, запълващи прякообектната позиция, например *water* („напоявам“), *milk* („доя“), *squeeze* („изстисквам“), *stroke* („галя“). Двайсет и един глагола, сред които *paint* („боядисвам“), *load* („товаря“) и др., са категоризирани като предикати, налагащи умерени селективни ограничения на аргумента си в същата синтактична позиция, а други двайсет и един – например *play* („играя“), *remove* („отстранявам, премахвам“) и др. – като неограничени по отношение на селективните ограничения върху аргумента в прякообектната позиция. След провеждането на тестовете глаголите, избрани за включване в експеримента, са групирани в три категории. Всеки глагол, за който 55% от предложените от носителите на езика отговори не надхвърлят три уникални запълващи фрази, е причислен към категорията на предикатите, налагащи най-строги селективни ограничения. В случай че уникалните отговори, представляващи минимум 55% от общия брой, са между четири и шест, предикатът се категоризира в групата на умерено ограничените

глаголи, а при седем или повече уникални фрази, съответстващи на минимум 55% от общия брой отговори, глаголът е отнесен към категорията на предикатите, налагащи най-малко ограничения върху прекия си обект. При преценката на уникалността на отговорите се вземат предвид само опорите на именните фрази, а предложните и адвербиалните фрази са изключени от изследването. Глаголите стимули са подбрани и с оглед на два други фактора: възможността аргументите им да бъдат представени чрез изображения, и наличието на подходящи фонологични кохорти. Включена е и неутрална категория глаголи като *choose*, *pick*, *select* („избирам, подбирам“) и *take* („вземам“), които не налагат селективни ограничения върху прекия си обект. За целите на експеримента са избрани по 16 глагола от всяка категория. Всяко изречение стимул се състои от пет думи: подлог [със семантичната роля АГЕНС], глагол, определителен член, съществително в прякообектна позиция и наречие. В подложната позиция са използвани следните пет лични имена: *Alex* („Алекс“), *Sam* („Сам“), *Charlie* („Чарли“), *Jesse* („Джеси“) и *Mary* („Мери“). Последната позиция на всяко изречение е заета от наречие, като по този начин се цели избягването на нежеланите ефекти, които биха могли да възникнат при попадането на тестваното съществително в изреченския край. 2,03% от данните – тестови данни с по-малко от 25% регистрирани точки на взирание (*gaze points*) – са изключени от по-нататъшния анализ. Преди началото на експеримента участниците попълниха онлайн тестове за проверка на степента на владеење на граматиката (https://www.examenglish.com/leveltest/grammar_level_test.htm) и лексиката (<http://vocabulary.ugent.be/>) на английския език. Резултатите от тестовете са включени в анализа като прогнозни променливи за представянето на участниците в последвалия експеримент. Статистическият анализ на събраните данни за зрителното поведение на тестваните лица включва дисперсионен анализ (ANOVA) за определяне на ефекта, оказван от категорията, към която е причислен съответният глагол, както и ковариационен анализ (ANCOVA) за установяване на ефекта на степента на владеење на езика (граматика и лексика) върху представянето на участниците.

Резултати

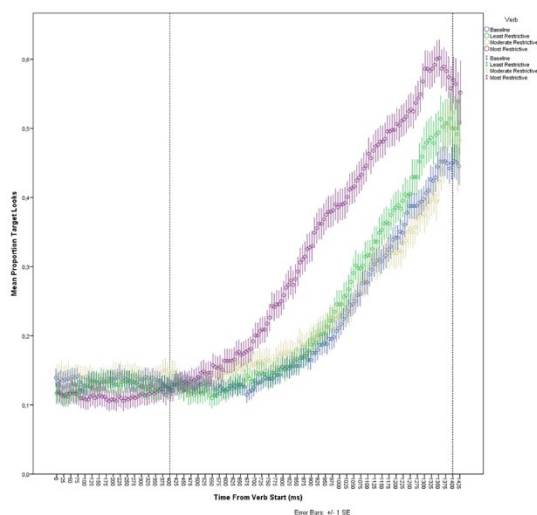
Изпреварващото насочване на погледа към целевия елемент е индикатор, че участникът разбира значението на глагола и притежава способността да го съотнесе с вероятна ситуация, описваща референтите на аргумента или аргументите му (Brock et al. 2008). Фигура 2 показва изменението на продължителността на задържане на погледа (*gaze proportions*) върху целевото изображение от началния момент на слуховото възприемане на глагола при групата с по-слабо владеење на английски език, а фигура 3 – стойностите на същия показател при групата на напредналите.



Фигура 2. Върху абсцисата са нанесени стойностите на съотношението на времето на задържане на погледа върху целевия елемент, а върху ординатата – времето (в милисекунди) спрямо началния момент на слуховото възприемане на глагола в изречението. VerbCat 0 = базовата неутрална категория глаголи, 1 = категорията, налагаща най-малки селективни ограничения, 2 = категорията, налагаща умерени селективни ограничения, 3 = категорията, налагаща най-строги селективни ограничения. Плътните вертикални линии, съответстващи на 400 и 1400 ms, обозначават началната и крайната точка на критичния времеви прозорец, а пунктираната вертикална линия – усредненото време, бележещо началото на слуховото възприемане на прекия обект.

При групата на участниците с по-ниска степен на владеене на езика средното начално време за насочване на погледа към целта е 651 ms, резултат, показателен за наличието на предварително очакване относно запълващия елемент в прякообектна позиция. Нарастването на продължителността на задържане на погледа върху целевото изображение, онагледена от фигура 2, показва по-голям ефект при глаголите, налагащи по-строги селективни ограничения върху прекия си обект, спрямо по-слабо ограничаващите глаголи, като този ефект продължава докъм 1400 ms след началото на слуховото възприемане на глагола. Извършена е пост-хок съпоставка по двойки на глаголните категории с цел да се установи дали помежду им съществува значима разлика в продължителността на задържане на погледа. Към стойностите на статистическа значимост е приложена корекция по Бонферони за сравнения по множество параметри. Продължителността на задържане на погледа за глаголите от най-ограничаващата по отношение на семантиката на обекта категория 3 е значително по-голяма спрямо стойностите на същия показател както за нулевата (базовата неутрална) категория, така и за категория 1 (категорията с най-малки се-

лективни ограничения за глаголния обект). Не се наблюдават значими разлики между глаголите от категория 2 и категория 3. С други думи, единствената категория, при която е регистрирано значително нарастване на задържането на погледа върху целевото изображение спрямо останалите категории, е групата на най-ограничаващите глаголи, и то само в съпоставка с най-малко ограничаващата категория 1 и с базовата категория.

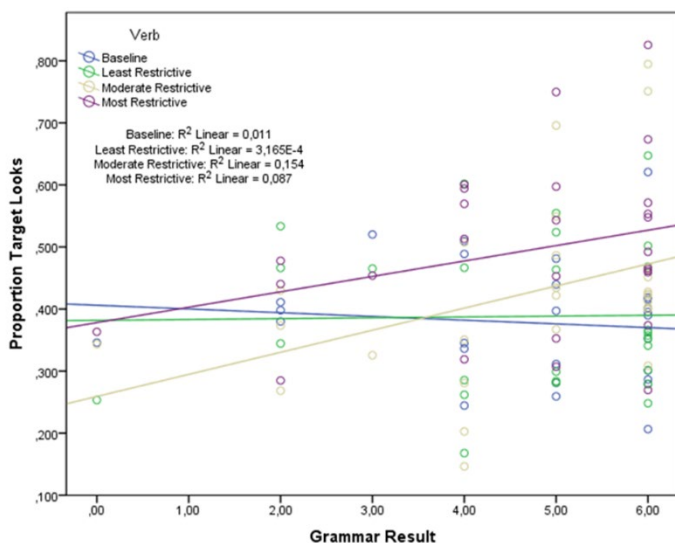


Фигура 3. Средно съотношение на задържането на погледа върху целевото изображение за глаголите от всички типове. Пунктираната линия обозначава критичния времеви прозорец.

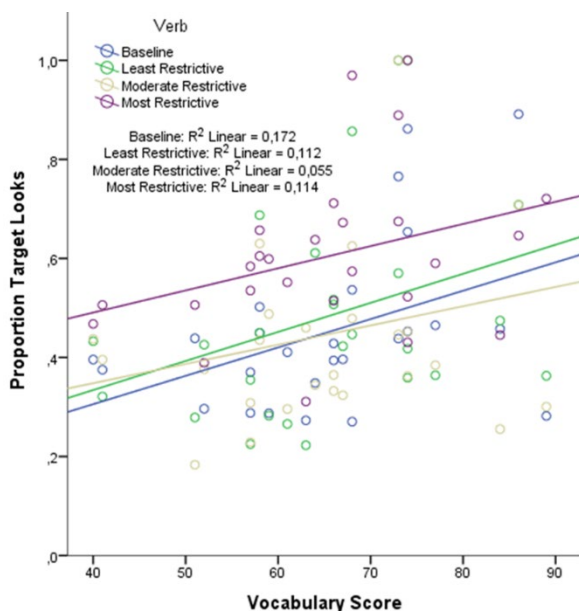
Постхок съпоставката по двойки на продължителността на задържане на погледа върху целевото изображение при групата на напредналите показва, че при глаголите от най-ограничената категория са регистрирани значително по-голям брой изпреварващи погледи, което предполага, че тези предикати се различават съществено от всички останали типове. Наред с това между другите три групи глаголи не са установени различия по отношение на очакванията за фразата, която ще се появи в позицията на прекия обект. Въпреки че тези резултати изглеждат близки до получените при участниците с по-ниска степен на владеене на английски като втори език, може да се открият и редица различия. При лицата от групата на по-напредналите насочването на погледа към целевото изображение започва около 600 ms след началото на слуховото възприемане на глагола, докато при групата на по-малко напредналите то се регистрира едва около 700 ms след този момент. Наред с това при най-ограничаващата категория предикати между двете групи участници се регистрира статистически значима разлика в средната продължителност на задържане на погледа върху целта в рамките на критичния времеви прозорец (400 – 1400 ms) – съответно

0,49112 за групата на владеещите езика в по-ниска степен и 0,60423 за групата на по-напредналите ($p < 0,01$).

Ковариационният анализ (ANCOVA) на резултатите при първата група показва статистически значим вътрешносубектен ефект от резултатите от граматичния тест – $F(3,66) = 3,202$, $p = 0,029$, $\eta^2 = 0,127$, но не и от резултатите от тестовете за проверка на лексикалните познания – $F(3,66) = 1,516$, $p = 0,218$, $\eta^2 = 0,064$. В рамките на същата група не е установен статистически значим междусубектен ефект (фигура 4). При групата на напредналите ковариационният анализ (ANCOVA) показва междусубектен ефект от степента на владеене на лексиката, който се доближава до статистически значими стойности ($F(3,871)$, $p = 0,61$, $\eta^2 = 0,139$) (фигура 5). Иначе казано, при лицата от групата на напредналите с по-високи резултати на лексикалните тестове се наблюдава по-ранно насочване на погледа и количествено повече погледи към целевото изображение независимо от глаголната категория, докато при групата с по-ниска степен на владеене на езика ефектът от глаголната категория е наблюдаем в зрителното поведение само на участниците, показали високи резултати на граматичните тестове.



Фигура 4. Взаимовръзка между съотношението на времето на задържане на погледа върху целта и резултатите от граматичните тестове при групата на по-малко напредналите. Тенденцията на нарастване на броя на погледите към целевите изображения при глаголите от категория 2 и категория 3 спрямо предикатите от базовата категория и категория 1 се наблюдава единствено при участниците, показали степен на владеене на граматиката, съответстваща на ниво на езикова компетентност C1 и C2.



Фигура 5. Взаимовръзка между съотношението на времето на задържане на погледа върху целта и резултатите от тестовете за владеење на лексиката при групата на напредналите. При участниците, показали високи резултатите на лексикалните тестове, се наблюдава по-продължително задържане на погледа върху целевите изображения независимо от категорията на глаголите, използвани в тестовите изречения.

Изследването показва, че изпреварващото зрително поведение се предопределя от различни езикови умения в зависимост от степента на владеење на втория език. При участниците от групата на по-напредналите лексикалното познание е предиктор за способността за предсказване на именната фраза, която е най-вероятно да се появи в позицията на вътрешния аргумент при всички глаголи независимо от категорията им, докато при изследваните лица от групата на по-малко напредналите по-добър предиктор за чувствителността към аргументната структура е граматичната компетентност, и то само при глаголите, налагащи най-строги селективни ограничения върху прекия си обект. Тези наблюдения насочват към заключението, че владеещите втори език на по-ниско ниво проявяват чувствителност предимно към глаголите с най-ограничени аргументни рамки, при които като приемлива реализация на даден аргумент се появяват малък брой именни фрази, докато в групата на напредналите чувствителността към аргументната структура – като индикатори на която служат изпреварващото зрително поведение и ефектът на лексикалното познание – се наблюдава при всички категории независимо от селективните ограничения, налагани на аргументите. Резултатите представят доказателства за нали-

чие на по-висока чувствителност към аргументната структура при лицата, владеещи втори език на по-високо ниво, за разлика от по-слабо изразеното изпреварващо поведение при лицата с по-ниска степен на владеене на езика, което при това се проявява само спрямо глаголите с най-строги селективни изисквания към вътрешния аргумент. Допълнително потвърждение за това различие се открива във времевите характеристики на зрителното поведение на участниците: при по-напредналите насочването на погледа към целевото изображение се регистрира в по-ранен времеви прозорец след началото на слуховото възприемане на глагола в сравнение с по-малко напредналите. При втората група се наблюдава и по-слабо отчетливо зрително поведение, както се вижда от по-малката разлика в наклона на правите, изобразяващи стойностите за най-ограничаващите (категория 3) и умерено ограничаващите (категория 2) глаголи.

4. Формулен език. Идиоми

Терминът *формулен език* (*formulaic language*) служи за назоваване на последователности от думи, които се срещат в речта с висока честота: *fish and chips, don't worry about it, take the wheel* („риба с пържени картофи“, „не си го слагай на сърце“, „поемам кормилото“). В общия случай това са използвани във всекидневното общуване несвободни фрази, характеризирани се с неизменяема граматична форма (Kuiper 2017). Според различни оценки този вид изрази съставляват между една трета и половината от използваната в ежедневието реч (Conklin, Schmitt 2012), като се смята, че употребата им улеснява езиковата обработка, като намалява натоварването върху работната памет. Съществуват предположения, че тези многокомпонентни цялости играят централна роля в общуването и еволюцията на езика, тъй като подпомагат преодоляването на затрудненията пред бързото и ефективно общуване, произтичащи от способността на човешкия мозък да обработва само постъпващата в момента информация (т.нар. *Now-or-Never bottleneck constraint*) (Christiansen, Chater 2016; Isbilen, Christiansen 2020). По същността си интерпретацията на тези изрази има некомпозиционен характер, т.е. цялостното им значение не може да се изведе непосредствено от значенията на съставляващите ги лексикални елементи. Формулният език включва множество различни категории, към които спадат и идиомите, метафорите, клишетата и фразеологичните единици. Дали и доколко интерпретацията на даден израз е небуквална, зависи от типа му. Поради спецификите си категорията на идиомите, които заемат централно място сред несвободните фрази, е обект на интерес от страна както на теорията за езика, така и на експерименталните изследвания.

В литературата са предлагани различни теории за природата на идиомите и механизмите на обработването им от мозъка. Главният въпрос, който си поставят представителите на отделните подходи, е дали достъпът до образното (небуквалното) значение на израза в менталния лексикон се

осъществява директно, по начина, по който се извличат значенията на думите, или процесът е предшестван от буквалната интерпретация на конституентите му. Интерес представлява и въпросът в какъв ред се извършват тези операции. Според един от подходите, основаващи се на хипотезата, че обработката на несвободните изрази протича на едно-единствено ниво, т.нар. модел на прекия достъп (*Direct Access model*), не е необходима предварителна активация на буквалното значение или значения (*Gibbs 1994*), а преносната интерпретация се извлича непосредствено с помощта на контекстови маркери (*cues*). За разлика от него стандартният прагматичен подход (*Standard Pragmatic approach*) постулира наличието на двустепенен процес, в чиято първа фаза се извършва активиране на буквалните значения, асоциирани с думите, съставлящи израза, а преносната небуквална интерпретация се конструира едва на следващия, втори етап (*Grice 1975*). Съществуват и подходи, според които двете нива на обработка протичат едновременно. Според един от тях хипотезата за лексикалната репрезентация (*Lexical Representation Hypothesis*; *Swinney, Cutler 1979*), идиомите се съхраняват като лексикални единици – факт, който позволява бързото им извличане от паметта – като паралелно с това протича и процес на лексикална декомпозиция. В съвременната изследователска парадигма се е наложило консенсусното виждане, че идиомите са комплексни изрази, чиято обработка не може да се сведе нито до процес на пряк лексикален достъп, нито до обработка на буквалното значение, а интерпретацията им зависи от редица фактори (*Vulchanova et al. 2015; Vulchanova et al. 2019a*). Например според конфигурационната хипотеза (*Configuration Hypothesis*) значението на идиомите е поделено между съставлящите ги единици, а самите те се обработват като сложни изрази, аналогично на други структури със сходна степен на синтактична сложност (*Sacciarì, Tabossi 1988*). При хибридният подход (*Titone, Connine 1999*) като най-важен критерий за различията при обработката на идиомите се извежда степента на разчлененост.

В поредица от неотдавнашни изследвания бе демонстрирано, че често в процеса на обработка на идиомите се активира буквалното им значение (*Vulchanova et al. 2019b; Chahboun et al. 2016; Milburn et al. 2021*), което в много случаи се конкурира с очакваната образна интерпретация. Дали буквалното тълкуване ще бъде „заобиколено“, зависи от разложимостта на израза, от индивидуалните различия между ползвателите на езика по отношение на това доколко идиомът им е познат (т.е. от обема на менталния лексикон), от етапа на развитие на дадено лице (дете или възрастен) и от нивото на владеење на езика, например дали за конкретния участник той е роден или чужд.

Въпреки изтъкваните тези и доказателства, че децата започват да изучават (буквалния) език под формата на цели сегменти от думи (*Vulchanova et al. 2011; Isbilen, Christiansen 2020*), усвояването на идиомите протича на

по-късен етап от развитието дори и в родния език. Това ги отличава от лексемите, които спадат към първите езикови единици, чието овладяване и използване започват още в пеленаческа възраст (Vulchanova, Vulchanov 2021). Това може да означава, че условията за успешно усвояване на думите и идиомите не са идентични. Наред с това в литературата е добре документирано наблюдението, че възрастните, изучаващи даден език като чужд, срещат трудности при овладяването на формулния език, употребяват по-малко формулни единици в сравнение с носителите на съответния втори език и е възможно да имат по-слабо развита способност да използват несвободните фрази, които са в състояние да идентифицират. Изглежда, най-голямото предизвикателство за усвояващите втори език е да преценят кога да прибегнат до обработката на дадена последователност от думи като единица с преносно значение. Това наблюдение се потвърждава и от редица изследвания, които показват, че скоростта, с която изучаващите даден език като чужд обработват идиомите и несрещани досега фрази с буквално значение, е сходна – факт, който на свой ред сочи, че двата типа изрази се обработват по аналогичен начин (Arnon, Christiansen 2017; Conklin, Schmitt 2008; Conklin, Schmitt 2012; Siyanova-Chanturia et al. 2011a; Siyanova-Chanturia et al. 2011b). Това контрастира с по-бързата обработка на идиомите, наблюдавана в родния език. Възниква въпросът кои са условията, които благоприятстват и улесняват обработката на идиоматичните изрази както в майчиния, така и в чуждия език. В хода на досегашните си изследвания сме идентифицирали два главни фактора, които имат отношение към този процес. Обикновено се смята, че наличието на подходящ контекст – само по себе си външно за идиомите условие – подпомага „конструирането“ на преносното тълкувание. Втори фактор, който може да повлияе върху това дали даден израз ще бъде обработен буквално, или според очакваната образна интерпертация, е честотата на срещане на отделните му компоненти в колокации и съвместно с останалите конституенти на израза. Това е вътрешноприсъщ за идиомите фактор, който по всяка вероятност подпомага разпознаването на този тип фрази и извличането им от паметта.

Степента, в която буквалното значение се конкурира с преносното, може да се предскаже въз основа на честотата на срещане на т.нар. идиоматичен „маркер“ (*idiomatic key*, думата, предопределяща идиоматичното значение) съвместно с останалите компоненти на идиома спрямо честотата му на срещане с други думи в последователности със свободно (неидиоматично) значение. При глаголите този подход може да се преформулира като вероятността дадена дума или израз да се появи като зависим от глагола конституент (негов пряк обект). Например глаголят *kick* („ритам“) най-често присъединява като свой комплемент именна фраза с опора *ball* („топка“), като според данните, получени от тестването на носители на английския език, вероятността за това е 54%, а според извадките от дос-

тъпни в интернет корпуси – 14,58%. Обратно, макар и да варира в различните корпуси, съвместната поява на *kick* и *bucket* (например в идиома *kick the bucket* – „ритам камбаната“), се характеризира с много ниска честота. Въз основа на това може да се предскаже, че тези два възможни компонента не се намират в конкуренция помежду си поради различната вероятност, с която всеки от тях може да се появи като обкръжение на глагола *kick* (теоретична дискусия по този въпрос е представена у Vulchanova et al. 2019a).

В поредица от провеждащи се в момента проучвания на проверка се подлага хипотезата дали носителите на английския език и лицата, които го изучават като чужд, разчитат на различни стратегии – по-конкретно, обработване „дума по дума“ и обработване на „цели фрази“ – при извличането на идиомите от паметта и възпроизвеждането им в речта. Предложеният експериментален дизайн си поставя за цел да се изследва приносът на двата главни фактора – наличието или отсъствието на влияещ на интерпретацията контекст и ролята на честотата на срещане на целевата дума от идиоматичния израз в колокации в съпоставка с фреквентността на високо- и нискочестотни думи в същото обкръжение. С оглед на това са подбрани идиоми, съдържащи глаголна опора и комплемент, което позволява контролиране на параметъра „честотата на срещане“ на единиците в комплементна позиция. Експерименталният дизайн се основава на парадигмата на зрителния свят (VWP) и се състои в проследяване и записване на погледите на участниците в проучването към региони от компютърния дисплей по време на слуховото възприемане на тестовите изречения. Зрителните стимули, които се появяват на екрана, представят възможни обекти на глагола: целевия преносен конституент, негов конкурент за същата позиция с висока честота на срещане в речта, конкурент за същата синтактична позиция с ниска честота на срещане и „уловка“ (*lure*). Слуховите стимули включват добре познати идиоматични изрази, употребени както в контекст, насочващ към преносното или към прякото им значение, така и самостоятелно (извън контекст). Участниците в експеримента са носители на английския език и лица, владеещи го във висока степен като чужд. Резултатите от наблюденията върху групата на участниците, за които езикът е нероден, са докладвани от Милбърн и колектив (Milburn et al. 2021), а описанието на комбинирания анализ на данните от двете групи в момента се подготвя за публикуване (Milburn et al., под печат). Получените данни подкрепят хипотезата, че носителите на езика и изучаващите го като чужд обработват идиомите по различен начин.

5. Заключителни бележки и изводи

Представеният обзор на съществуващите експериментални доказателства сочи, че чувствителността към особеностите на езиковата структура както при родния, така и при чуждия език се развива с течение на времето.

В статията бе изложено и онагледено виждането, че чувствителността към аргументната структура на глаголите е главен компонент както от родно-езиковата, така и от чуждоезиковата компетентност. При все че много от механизмите, на които се основава овладяването на двата езика, може да са едни и същи (например фонологичната и работната памет, усвояването, базиращо се на извличането на статистически закономерности (*statistical learning*, вж. Vulchanova, Vulchanov 2021), съществуват и редица различия. Обработката на езика зависи от индивидуалните характеристики на обучаемия, както и от това дали компетентността му е родно- или чуждоезиковата. Наред с това има изобилие от доказателства, че децата (при които все още тече процес на усвояване) и възрастните обработват езика по различен начин. Децата проявяват склонност към буквална интерпретация на идиомите, каквато при възрастните в общия случай се наблюдава по-рядко (Vulchanova et al 2011; Chahboun et al. 2016).

С течение на времето обработката става все по-автоматизирана и възрастните носители на езика по-често се осланят на усвояването му на по-големи сегменти, най-вероятно като стратегия за преодоляването на затрудненията в общуването (Christiansen, Chater 2021b; Milburn et al., под печат). Не на последно място, в статията се демонстрира, че идиомите предлагат ползотворен инструмент за изследване на чувствителността към лексикалните свойства на думите и механизмите на езиковата обработка и може да послужат като източник на надеждни наблюдения относно особеностите на буквалния език.

Благодарности

Настоящият текст е към научната програма на проекта „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“, подкрепен от Фонд „Научни изследвания“ по програма „Финансиране на фундаментални научни изследвания – 2020 г.“, Договор No КП-06-Н50/1 от 2020 г.

Превод от английски: *Светлозара Лесева*

REFERENCES

- Alloppenna et al. 1998: *Alloppenna, P. D., J. S. Magnuson, M. K., Tanenhaus*. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. – *Journal of Memory and Language*, 38(4), pp. 419–439. doi:10.1006/jmla.1997.2558.
- Altmann 1997: *Altmann, G. T. M.* Words and how we (eventually) find them. – *The ascent of Babel: an exploration of language, mind, and understanding*. Oxford: Oxford University Press, pp. 1–32.
- Altmann, Kamide 1999: *Altmann, G. T. M., Y. Kamide*. Incremental interpretation at verbs: Restricting the domain of subsequent reference. – *Cognition: International*

- nal Journal of Cognitive Science*, 73(3), pp. 247–264. doi:10.1016/S0010-0277(99)00059-1.
- Arnon, Christiansen 2017: *Arnon, I., M. Christiansen*. The Role of Multiword Building Blocks in Explaining L1–L2 Differences. – *Topics in Cognitive Science*, 9(3), pp. 621–636.
- Brock et al. 2008: *Brock, J., C. Norbury, S. Einav, K. Nation*. Do individuals with autism process words in context? Evidence from language-mediated eye movements. – *Cognition: International Journal of Cognitive Science*, 108(3), pp. 896–904. doi:10.1016/j.cognition.2008.06.007.
- Cabrera, Zubizarreta 2005: *Cabrera, M., M. L. Zubizarreta*. Overgeneralization of causatives and transfer in L2 Spanish and L2 English. – In: Eddington, D. (ed.). *Selected Proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*. Somerville, MA: Cascadilla Proceedings Project, pp. 15–30.
- Cacciari, Tabossi 1988: *Cacciari, C., P. Tabossi*. The comprehension of idioms. – *Journal of Memory and Language*, 27(6), pp. 668–683. [https://doi.org/10.1016/0749-596X\(88\)90014-9](https://doi.org/10.1016/0749-596X(88)90014-9).
- Chahboun et al. 2016: *Chahboun, S., V. Vulchanov, D. Saldaña, H. Eshuis, M. Vulchanova*. Can You Play with Fire and Not Hurt Yourself? A Comparative Study in Figurative Language Comprehension between Individuals with and without Autism Spectrum Disorder. – *PLoS ONE*, 11(12), e0168571. <https://doi.org/10.1371/journal.pone.0168571>.
- Christiansen, Chater 2016: *Christiansen, M. H., N. Chater*. The Now-or-Never bottleneck: A fundamental constraint on language. – *Behavioral and Brain Sciences*, 39, e62.
- Conklin, Schmitt 2008: *Conklin, K., N. Schmitt*. Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? – *Applied linguistics*, 29(1), pp. 72–89.
- Conklin, Schmitt 2012: Conklin, K., N. Schmitt. The Processing of Formulaic Language. – *Annual Review of Applied Linguistics*, 32, pp. 45–61.
- Dimitrova-Vulchanova 1999: Dimitrova-Vulchanova, M. *Verb semantics, diathesis and aspect*. Munich: Lincom Europa.
- Dimitrova-Vulchanova, Weisgerber 2007: *Dimitrova-Vulchanova, M. M. Weisgerber*. 'Concept,' 'Context' and Beyond: Manner of Motion Verbs, Encoding and Relevant Bits of Information. – *Southern Journal of Linguistics*, 29, pp. 52–76.
- Gibbs 1994: *Gibbs, R. The poetics of mind: Figurative thought, language and understanding*. Cambridge: Cambridge University Press.
- Grice 1975: *Grice, H. P.* Logic and conversation. – In: Cole, P., J. L. Morgan (eds.). *Syntax and Semantics 3: Speech Acts*. Cambridge, MA: Academic Press, pp. 41–58.
- Isbilen, Christiansen 2020: *Isbilen, E. S., M. H. Christiansen*. Chunk-based memory constraints on the cultural evolution of language. – *Topics in Cognitive Science*, 12, pp. 713–726.
- Jackendoff 2002: *Jackendoff, R.* What's in the lexicon? – In: Nooteboom, S., F. Weerman, F. Wijnen (eds.). *Storage and computation in the language faculty*. Dordrecht: Kluwer, pp. 23–58.

- Johnsen 2016: Johnsen, R. *Sensitivity to second language argument structure: An experimental study with Norwegian learners of English*. MA thesis, NTNU.
- Koenig et al. 2003: Koenig, J. P., G. Mauner, B. Bienvenue. Arguments for adjuncts. – *Cognition*, 89 (2), pp. 67–103.
- Kuiper 2017: Kuiper, K. *Formulaic Language*. *Oxford Bibliographies. Linguistics*. 10.1093/obo/9780199772810-0212.
- Larrañaga et al. 2012: Larrañaga, P., J. Treffers–Daller, F. Tidball, M.–C. Gil Ortega. L1 transfer in the acquisition of manner and path in Spanish by native speakers of English. – *International Journal of Bilingualism*, 16, pp. 117–138.
- Levelt 1993: Levelt, W. J. M. Accessing words in speech production: Stages, processes and representations. – In: Levelt, W. J. M. (ed.). *Lexical access in speech production*. Cambridge, MA: Blackwell Publishers, pp. 1–22.
- Milburn et al. 2021: Milburn, E., M. Vulchanova, V. Vulchanov. Collocational frequency and context effects on idiom processing in advanced L2 speakers. – *Canadian Journal of Experimental Psychology (Revue Canadienne De Psychologie Expérimentale)*. DOI: 10.1037/cep0000214.
- Milburn et al. (in preparation): Milburn, E., V. Vulchanov, J. Magnuson, M. Vulchanova. Native speakers kick buckets but learners kick doors: a comparison of native and non-native idiom comprehension.
- Perfetti 2007: Perfetti, C. Reading ability: Lexical quality to comprehension. – *Scientific Studies of Reading*, 11(4), pp. 357–383. doi:10.1080/10888430701530730.
- Perfetti, Hart 2002: Perfetti, C. A., L. Hart. The lexical quality hypothesis. – In: Verhoeven, L., C. Elbro, P. Reitsma (eds.). *Precursors of functional literacy*. Philadelphia, PA: John Benjamins, pp. 189–214.
- Pinker, Jackendoff 2005: Pinker, S., R. Jackendoff. The faculty of language: what’s special about it? – *Cognition*, 95(2), pp. 201–236. doi: 10.1016/j.cognition.2004.08.004.
- Reine 2016: Reine, A. *L2 speakers’ argument structure sensitivity inferring from the participant’s eye gazes: An eye-tracking experimental study with adult Norwegian learners of English*. MA thesis, NTNU.
- Sivanova-Chanturia et al. 2011a: Sivanova-Chanturia, A., K. Conklin, W. J. B. van Heuven. Seeing a phrase “time and again” matters: The role of phrasal frequency in the processing of multiword sequences. – *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(3), pp. 776–784.
- Sivanova-Chanturia et al. 2011b: Sivanova-Chanturia, A., K. Conklin, N. Schmitt. Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. – *Second Language Research*, 27(2), pp. 251–272.
- Stringer 2010: Stringer, D. The gloss trap. – In: Han, Z.-H., T. Cadierno (eds.). *Linguistic relativity in SLA: Thinking for speaking*. Clevedon, UK: Multilingual Matters, pp. 102–124. doi:10.21832/9781847692788-007.
- Swinney, Cutler 1979: Swinney, D. A., A. Cutler. The access and processing of idiomatic expressions. – *Journal of Verbal Learning and Verbal Behavior*, 18(5), pp. 523–534.

- Titone, Connine 1999: *Titone, D. A., C. M. Connine*. On the compositional and non-compositional nature of idiomatic expressions. – *Journal of Pragmatics*, 31(12), pp. 1655–1674.
- Treffers-Daller, Calude 2015: *Treffers-Daller, J., A. Calude*. The role of statistical learning in the acquisition of motion event construal in a second language. – *International Journal of Bilingual Education and Bilingualism*, 18 (5), pp. 602–623.
- Treffers-Daller, Tidball 2015: *Treffers-Daller, J., F. Tidball*. Can L2 learners learn new ways to conceptualise events? Evidence from motion event construal among English-speaking learners of French. – In: Schmitz, K., P. Gujarro-Fuentes, N. Mueller (eds.). *The acquisition of French in its different constellations*. Multilingual Matters.
- Vulchanova et al. 2011: *Vulchanova, M., V. Vulchanov, M. Stankova*. Idiom comprehension in the first language: A developmental study. – *Vigo International Journal of Applied Linguistics*, 8.
- Vulchanova et al. 2015: *Vulchanova, M., D. Saldaña, S. Chahboun, V. Vulchanov*. Figurative language processing in atypical populations: the ASD perspective. – *Frontiers in Human Neuroscience*, 9(24), pp. 1–11.
- Vulchanova et al. 2019a: *Vulchanova, M., E. Milburn, V. Vulchanov*. Boon or Burden? The Role of Compositional Meaning in Figurative Language Processing and Acquisition. – *Journal of Logic, Language and Information*, 28, pp. 359–387. <https://doi.org/10.1007/s10849-019-09282-7>.
- Vulchanova et al. 2019b: *Vulchanova, M., S. Chahboun, B. Galindo-Prieto, V. Vulchanov*. Gaze and motor traces of language processing: Evidence from autism spectrum disorders in comparison to typical controls. – *Cognitive Neuropsychology*, 36(7–8), pp. 383–409. <https://doi.org/10.1080/02643294.2019.1652155>.
- Vulchanova, Vulchanov 2021: *Vulchanova, M., V. Vulchanov*. Language development in typically developing children and children with developmental deficits: A concise qualitative review. – In: *Papers of the Institute for Bulgarian Language “Prof. Lyubomir Andreychin”*, 34, pp. 257–273. DOI: 10.7546/PIBL.XXXIV.21.10

✉ Проф. д-р Мила Вълчанова

✉ Д-р Валентин Вълчанов

Норвежки университет за науки и технологии
Edvard Bulls veg 1, Trondheim 7491, Норвегия

✉ Prof. Mila Vulchanova, PhD

✉ Valentin Vulchanov, PhD

Norwegian University of Science and Technology
Edvard Bulls veg 1, Trondheim 7491, Norway