Article:

Divjak, D. (2015) Exploring the grammar of perception A case study using data from Russian. Functions of Language, 22 (1). 44 - 68. ISSN 0929-998X

https://doi.org/10.1075/fol.22.1.03div

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

EXPLORING THE GRAMMAR OF PERCEPTION

A CASE STUDY USING DATA FROM RUSSIAN

Dagmar Divjak **d.divjak@shefield.ac.uk**

2015. Functions of Language 22 (1): 44-68
Sensory Perceptions in Language and Cognition
Edited by Rosario Caballero and Carita Paradis

Abstract

In this paper, I pursue the distributional hypothesis that the meaning of a word is derived from the linguistic contexts in which it occurs and apply it to verbs of perception. Differently from NLP implementations of the distributional hypothesis, I explicitly limit the range of variables to the grammatical domain and chart the way in which verbs of Vision, Hearing and Touch are used, morphologically and syntactically, in a representative sample of corpus data. Some aspects of experience are so central and pervasive that reference to them has grammaticalized (Divjak 2010; see also Newman 2008; Janda & Lyashevskaya 2011).

The aim is, firstly, to determine to which extent a verb's grammatical context alone allows us to classify utterances according to the perception type, and, secondly, to chart the similarities and differences in the verbs' preference for morphological markers and syntactic constructions. If contexts are highly specialized, language structure, as it is witnessed in use, could assist sensory impaired speakers in building up viable representations of concepts, even if sensory experience is lacking. If, in addition, similarities between certain sensory perception verbs are high, sensory impaired speakers could use these similarities to perform analogical mapping across senses and ground concepts relating to the impaired sense in a cognate sensory experience.

The findings are relevant for concept acquisition and representation in general and for concept acquisition and representation in sensory impaired populations, such as the blind, in particular.

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

1. Background

The human capacity for acquiring and representing meaning has long puzzled philosophers, linguists, psychologists and cognitive scientists alike. Theories for understanding the nature and content of concepts cover a spectrum from innate to empiricist and from purely mentalist to fully embodied. Although empirical approaches to conceptual content acknowledge that concepts are made up of information gathered through our bodies and senses as well as through language, research has foregrounded our modal, sensory-motor experiences, at the expense of our linguistic experience, which is typically considered a-modal within the Cognitive Sciences. After a brief introduction to the building blocks of concepts (Section 1.1), I will present the Distributional Hypothesis (Section 1.2) as a way of capturing the contribution language makes to concept formation.

1.1 Concepts, and what they are made of

Concepts, the constituents of thoughts, can be defined, at the most general level, as mental structures corresponding to a specific entity or class of entities, be they concrete or abstract (Matthews 2007). In other words, concepts generalize over experiences. Three sources of information for concept formation are commonly acknowledged, i.e. direct sensory-motor experience (experiencing yourself), indirect sensory-motor experience (witnessing others experience), as well as experience with language.

In recent years, the debate about innate concepts has been reinvigorated as advances in cognitive science have provided new tools for revisiting the dispute (Wilson 2002). Evidence has accumulated that concepts can be acquired on the basis of experience using a few relatively simple general-purpose cognitive mechanisms. The discussion is therefore shifting to the question of which experiences feed into the mix and how much of sensory-motor experience is retained in the mental representation. Recently, Vigliocco et al. (2009) have proposed an account of semantic representations that recognizes affective and linguistic experience in addition to sensory-motor information.

With this turn, the question of how meaning is learned from statistical distribution across texts is again gaining popularity in cognitive circles. In this study, I focus on the question of what language has to offer the learners in their quest for the meaning of lexemes and the concepts they give access to.

1.2 Distributional learning

The distributional hypothesis (Harris 1954) driving this research question states that the meaning of a word is derived from the linguistic contexts in which it occurs. Words that are more similar in meaning should occur in a larger number of shared contexts.

Existing implementations of the distributional hypothesis are found within research on natural language processing. A range of models has been designed that rely on the distributional hypothesis, with Vector Space Models such as Latent Semantic Analysis (Landauer et al. 1998) and Probabilistic Topic Models (Hofmann 1999) as the best known types in linguistic circles. “Context” is a crucial concept for these models, and it is the property most accounts diverge on. Issues typically discussed in conjunction with “context” relate to its definition and structuring.

The size of the contextual unit taken into consideration in NLP applications varies from two-word windows to an entire text. It has been found that larger contexts yield topical information suited for information retrieval tasks. Smaller contexts, on the other hand, reveal knowledge akin to lexical semantic competence. The structure of the context is typically not taken into account, which has given these implementations the name of “Bag of words”-models. The reason for this is that adding structural information to text requires a non-negligible amount of pre-processing and easily suffers from data sparseness, while the empirical evidence for the supremacy of refined contexts remains scarce (Sahlgren 2008: 47-48).

I use the Behavioural Profiling (BP) implementation of the distributional hypothesis (as presented in Divjak 2003; Divjak 2006; Divjak & Gries 2006; Divjak 2010; see references therein to related work by Gries, Arppe, Janda and collaborators). From the point of view of a linguist, BPs improve on the NLP implementations by narrowing down the context window to a "natural" unit of expression, i.e. a sentence or clause, and by annotating the entire context manually for a range of morphosyntactic, syntactic or semantic properties.

Divjak (2006) argues that constructional networks (as used by Apresjan 1967; Levin 1993 and others to delineate semantically similar groups of verbs) reveal coarse-grained meaning similarities and differences between verbs because constructions outline the meaning contours of the verbs that occur in them (different from lexical elements that convey detailed information; see also Li and Brew (2008). Janda & Solovyev (2009) show that one can rely on constructions to distinguish between semantically similar nouns in Russian.

At a lower level of granularity, Šteinfeldt (1970) observed that Russian verbs vary in the frequency distribution of their paradigm forms and Karlsson (1986: 27) concluded on the basis of Finnish data that meaning properties are reflected in the use of forms. In a medium-scale practical application of this idea, Janda & Lyashevskaya (2011) tracked preferences in tense, aspect mood marking (TAM) to delineate semantically coherent subgroups of verbs. Divjak (2010) showed that TAM markings on verbs are the variables that distinguish best between near-synonymous verbs.

The fact that differences in the morphosyntactic and syntactic distribution of verbs are connected to the verbs' semantics provides learners with a more powerful bootstrapping device for acquiring lexical meaning (cf. Landau & Gleitman 1985); if the meaning of a lexeme is not exclusively accessed through the meaning of other lexemes but is accessible through morphosyntax and syntax as well, infinite regress is avoided. For this reason, I will limit myself to a BP of morphosyntactic and syntactic properties in this paper, leaving semantic properties for future investigation.

Of course, representations of meanings of words induced from analysis of textual data on morphosyntactic and syntactic properties alone are without doubt "bloodless and sterile" (Landauer et al. 1998). There is however growing evidence that much sensory and bodily experience is encoded in language (Louwerse & Jeuniaux 2010). Language users can thus rely on the linguistic system as a shortcut to the perceptual system (Louwerse 2011). Some go as far

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

as to claim that perceptual and distributional streams of data are redundant streams and speakers can attend to either type (Riordan & Jones 2010).

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

2. Data & Method

I analyse usage data of Russian verbs of seeing, hearing and touching to determine whether grammar distinguishes clearly between these verbs while allowing them to share a common constructional base that would facilitate drawing conceptual and experiential analogies.

2.1 Verb selection

In English, it seems straightforward to select the members of the agentive/experiential Vision, Touch and Hearing pairs, i.e. look/see for Vision, touch/feel for Touch and listen/hear for Hearing. In Russian, selecting the basic verbs in each category and determining the most neutral perfective counterparts is a task riddled with difficulty, in particular for the domain of Touch. The relation between touch and feel in English and between the equivalents for these verbs in Russian is not identical to the relation we find in look/see and listen/hear and their Russian equivalents: Feel covers a much broader domain outside its primary area of tactile perception than see and hear do. A detailed analysis of the differences remains outside the scope of this study.

In Russian, perception verbs seem to form a network, rather than pairs, both lexically and aspectually. I settled for the three "main"[1] senses (Vision, Hearing, Touch) as expressed by means of the most frequent verbs (see Table 2 below) that participate in a voice-like opposition (Agent/Experiencer, i.e. look vs see, listen vs hear, touch vs feel) and an aspectual opposition (imperfective/perfective), yielding the 3 pairs presented in Table (1).

Table 1. The six perception verb pairs in Russian

Perception type	Agentive perception		Experiential perception	
	Imperfective	Perfective	Imperfective	Perfective
VISION	smotret'	posmotret'	videt'	uvidet'
HEARING	slušat'	poslušat'	slyšat'	uslyšat'

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

TOUCH	trogat’	potrogat’/tronut’	čuvstvovat’	počuvstvovat’

According to Nesset et al. (2008), agentive verbs prefer **po-** prefixation while their experiential counterparts attract **u-**. This generalization holds for Vision that contrasts agentive **po***smotret’* with experiential **u***videt’* and Hearing where we observe agentive **po***slušat’* but experiential **u***slyšat’*. The domain of Touch behaves differently as it can employ po- for perfectivizing both agentive and experiential perception, i.e. **po**trog*at’* vs **po***čuvstvovat’*, while also offering the more frequently used perfective *tronut’*.[2]

In what follows, I provide an example for each verb; first of the agentive verb and then of the experiential one.

VISION:

(1) Я долго не решался и *посмотрев* на небо понял — стоит попробовать. [Женщина + мужчина: Психология любви (форум) (2004)]

I took a long time to decide but having looked at the sky I understood – it is worth a try.

(2) Николай повернулся и побрел в комнату, догадываясь уже, кого он там *увидит*. [Евгений Лукин. Delirium tremens (Страсти по Николаю) (1997)]

Nikolaj turned around and wandered off into the room, already guessing who he’d see there.

HEARING:

(3) В магазинах, ларьках и на рынках никто не желает специально для меня срывать целлофан с компакт-диска и ставить CD на плейер, чтобы я чего-то там *послушал*. [В. А. Александр. Музыка через трубочку (1997) // «Столица», 1997.06.10]

In the shops, stands and on markets no one is particularly keen to rip the plastic off a CD for me and to put the CD in the player, so that I could listen to it there.

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

(4) Именно там меня *услышал* дирижер Даниэль Баренбойм, с которым я потом пела в Берлине. [Русская дива (2003) // «Лебедь» (Бостон), 2003.11.01]

It's there that director Daniel Barenboim heard me, with whom I later sang in Berlin.

TOUCH:

(5) Он нежно *тронул* жену за плечо и добавил грудным голосом: [Леонид Юзефович. Дом свиданий (2001)]

He tenderly touched his wife on the shoulder and added in a deep voice: (…)

(6) До 32000 человек […] *почувствовали* в пик морозов недостаток тепла в квартирах. [Горячая тема. Мороз ломает зубы о марийские теплотрассы (2003) // «Марийская правда» (Йошкар-Ола), 2003.01.10]

Up to 32 00 people felt, at the peak of the cold spell, the lack of warmth in their apartments.

There are, however, many more verbs available to express Touch; these are listed below with their English translation (taken from the Oxford Russian-English dictionary) but will remain outside the scope of this article because they do not categorize the event at the basic level of categorization. Apart from *čuvstvovat'/počuvstvovat'* ('feel, sense'), illustrated in (6), which is not touch specific and can also be used for smell[4], experiential feeling can be encoded with *oščuščat'/oščutit'* ('feel, sense, experience'), *osjazat'* ('feel')[3], and vulgar *čujat'/počujat'* ('scent, smell'; (fig) 'sense, feel'). Agentive touching is not only rendered by *trogat'/potrogat'* or *tronut'* ('touch') as illustrated in (5), but also by *(pri)kasat'sja/(pri)kosnut'sja* ('touch lightly'), *ščupat'/poščupat'* ('feel for, touch, probe'), *šarit'* ('grope about, feel, fumble') as well as *probovat'/poprobovat'* ('test'; again, not touch specific). All of these verbs combine with a range of prefixes to form perfectives.

2.2 Corpus data

Data on the 6 most general verb pairs were extracted from the Russian National Corpus (RNA). A contemporary 78-million-word written subcorpus was set, spanning the period 1992-2012; the entire 10-million-word oral subcoprus was used to ensure a large enough number of observations. The subcorpora were searched for each lexical item used as "Verb".

An overview of the overall frequencies of occurrence is given in Table (2) that shows frequency of mention supremacy of Vision over Hearing, followed by Touch; of experiential perception over agentive perception; and of imperfective aspect over perfective aspect within written data. In the oral corpus, a similar trend is identified, except for agentive listening versus experiential hearing that seem to be equally frequent.

Table 2. Raw frequencies per verb in the RNC

Perception	**Verb**	**Type**	**Aspect**	**# in written subcorpus**	**# in oral corpus**
Vision	Videt'	**Experiential**	**Impf**	134,004	30,382
	Uvidet'	**Experiential**	Pf	61,746	5302
	Smotret'	Agentive	**Impf**	95,540	20,204
	Posmotret'	Agentive	Pf	42,318	12,400
Hearing	Slyšat'	**Experiential**	**Impf**	34,520	13,480
	Uslyšat'	**Experiential**	Pf	22,526	1768
	Slušat'	Agentive	**Impf**	29,696	13,744
	Poslušat'	Agentive	Pf	6542	3512
Touch	Čuvstvovat'	**Experiential**	**Impf**	33,730	3742

	Počuvstvovat'	**Experiential**	Pf	16,638	910
	Trogat'	Agentive	**Impf**	4588	1466
	Tronut'	Agentive	Pf	3854	488

The first 1000 examples per lexeme were downloaded and the samples were further cleaned to leave 1 per author, which ensures independence of observations; finally, the cleaned samples were randomized and the first 300 examples per perception type were selected. Seven observations were excluded,[5] yielding a dataset of 893 observations in total.

2.3 Annotation

As explained in Section 1.2, it is the aim of this paper to explore to which extent the grammar-as-bootstrapping device would aid in acquiring verbs of perception, i.e. to establish whether language distinguishes grammatically between Vision, Hearing and Touch (and not only between look and see, which Landau and Gleitman (1985) established for English) and to find out where the parallels are and where the differences lie. For this reason the extractions were annotated for so-called "skeletal" information only (Divjak 2006), as summarized in Table (3).

Table 3. Variables and variable levels included in the sample annotation

	Variable (levels)	**Variable level label**
Verb	aspect (2)	(imperfective, perfective)
	mode (6)	(infinitive, indicative, imperative, conditional, gerund, participle[6])
	tense (5)	(past, present, future, non-past, none[7])
	number (3)	(singular, plural, none)
	person (4)	(1, 2, 3, none)

	voice (2)	(agentive, experiential)
Construction	polarity (2)	(negative, positive)
	argument structure types (9)	(lexeme as adjective, lexeme as adverb, verb only, verb followed by adjective/adverb, verb followed by noun, verb followed by direct object, verb followed by preposition, verb followed by that-clause, verb followed by adverbial clause) [8]
Situation	situation (2)	(spoken, written)

Most of the variables and variable levels listed in Table (3), except for Construction, need no further explanation. Examples of each of the constructions identified in the table under "argument structure type" are given below in (7) through (15).

(7) [transitive use of verb: verb with direct object]

Для того чтобы *послушать птичье пение*, нужно поселиться в окрестностях поселка Морское, например в отеле «Досуг», расположенном на берегу залива. [Янтарный край России (2000) // «Туризм и образование», 2000.06.15]

To hear bird song, you have to settle in the vicinity of the settlement Morskoe, for example in hotel "Leisure" that is situated on the bank of the bay.

(8) [intransitive use of verb: verb only]

Охлобвстин [sic], муж Качанов снял вот в сентябре / всех приглашаю / кто *слышит и видит_*/ так вот / [Беседа Д. Диброва с И. Охлобыстиным в эфире телепередачи «Антропология», НТВ // Архив Хельсинкского университета, 1999]

[…] I invite everyone who can hear and see […]

(9) [verb followed by adverbial clause]

Я вчера в книжном магазине *видела, как* одна женщина покупала своему ребенку учебники за 4-5 класс, хотя он учится во втором. [Надежда Куртинина, Марина Булгакова. Кризис глазами детей (1998) // «Нижегородские губернские ведомости», 1998.09.11]

I saw yesterday in the bookstore how one lady bought her child a textbook for 4th or 5th grade, although he's only in 2nd grade.

(10) [verb used as adverbial participle]

С ней пела московская 17 — летняя молодежь, которая, *видимо*, от мам и пап в хромосомном наборе получила тягу к ритмам в стиле «Ра-ра-распутин… ». [Парк культуры (1997) // «Столица», 1997.06.17]

With her sang the 17-year old Moscow youth, who clearly received in their chromosome set from mom and dad an inclination for rhythms of the style "ra-ra Rasputin …"

(11) [verb followed by pronoun "self" and noun (in instrumental)]

Для молодежи очень важно не *чувствовать себя неудачниками*, никто не хочет быть таким как он, и они начинают задумываться о том, что для этого надо делать. [Елена Елкина. Эпика 2001 (2001) // «Рекламный мир», 2001.12.25]

It is very important for the youth not to feel (like) failures (…)

(12) [verb followed by pronoun "self" and adjective (in instrumental) or by adverb]

Внимание Франсуа Миттерана было столь же острым, но в нем часто была ирония, дистанция, так что ты *чувствовал себя* одновременно *польщенным и задетым*… [Екатерина Демьянова, Алексей Каменский. Миттеран и его эпоха (1996) // «Коммерсантъ-Daily», 1996.01.20]

The attention of Mitterand was so sharp, but in it there was often irony and distance, so that you felt at the same time flattered and offended_…

(13) [verb used as adjectival participle]

Мужчина должен быть мыслящим и *чувствующим человеком*, ощущающим себя неуютно, если где-то происходит несправедливость. [Ольга Митина и др. Идеология маскулинности в России: постановка проблемы и экспериментальное исследование (2003) // «Общественные науки и современность», 2003.04.30]

A man has to be a thinking and feeling person, who feels uneasy if something unfair happens.

(14) [verb and prepositional phrase]

Потом мы долго лежали и молча *смотрели на* свечи, на то, как свет пламени играет на коже. [Ольга Зуева. Нас любовь кружила до утра // «Даша», 2003]

Afterwards we lay there for a long time and looked at the candles, at how the light plays with flames on the skin.

(15) [verb followed by that-clause]

Когда ты скажешь эту фразу 10-й раз, *почувствуешь, что* говоришь правду. [Егор Михалков-Кончаловский: Гуталин и пылесос как средства борьбы со стрессом (1997) // «Здоровье», 1997.12.15]

When you'll have said this phrase the 10th time, you'll feel that you are saying the truth.

In addition to the structural variables, one coarse meaning-related variable "reading" (literal, non-literal) was coded for: a tag "literal" was assigned to an extraction if the situation represented required activation of sensory perceptors, such as (16). If sensory perceptors did not need to be activated, the tag “non-literal” was assigned, as in (17) (see Paradis & Eeg-Olofsson 2013 for a discussion of metaphorization within the domain of perception).[9]

(16) Ее приятно трогать руками, она хранит в себе тепло желтого солнечного света, впитывает наше дыхание. [Андалусские мотивы (2001) // «Ландшафтный дизайн», 2001.03.15]

It is nice to touch her with your hands, she keeps in herself the warmth of the yellow sun light, absorbs our breathing.

(17) Пьеса окончена, а кажется, что ее мелодия, словно тронувшая сердце струна, продолжает звенеть. [Добрый молодец в министерском кресле. Обратная связь (2002) // «Известия», 2002.07.02]

The play is over, but it seems that its melody, like a string that has touched the heart, continues to sound.

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

3. Results

A first aim of this study (Section 3.1) was to identify clear similarities and striking differences between Vision and the two other senses studied that could aid (sensory impaired) speakers in building up an overarching concept of seeing, as suggested by Landau & Gleitman (1985) for individual verbs of seeing. A second goal is to determine how well perception type can be predicted from morphological and syntactic variables alone (Section 3.2).

3.1. Univariate analysis

In order to identify significant attractions between variables and outcomes, all variables were initially explored individually with respect to one of three outcomes: The three perception types (Section 3.1.1), the agentive vs experiential modes of perception (Section 3.1.2) and the members of each pair of verbs (Section 3.1.3). Given that independence of observations had been assured at the data collection stage, a simple Chi-squared test could be run on candidate variables to establish whether there is significant dependence between the variable and the outcome. Due to the number of tests run (one for each variable), the conservative Bonferroni correction was applied and the significance level was lowered from 0.05 to 0.005 throughout this section to maintain an overall alpha level of 0.05. Standardized Pearson residuals were inspected to find the cells that make the largest contribution to the Chi-squared values and to identify the direction of any deviations (positive and thus overrepresented or negative or underrepresented) from the expected values. The largest significant deviance from expectation is reported; complete tables are available in the on-line Appendix.

The results reported in this section do not imply that the properties singled out would occur together; rather, these properties are identified as individually distinctive for a perception type or verb. Properties that are not listed are distributed in a way that does not form a significant deviation from what would be expected, given the marginal values, if there were no association between the two variables investigated. So, although there is a clear and specific core for each type, there are many similarities among the types too. Section 3.2. will reveal how these properties interact to form the contexts that are typical for each verb.

3.1.1 Perception types

This analysis takes as input the rows in column 1 of Table (1) and compares the behaviour of perception types, aggregating findings over the individual verbs within each perception type.

For Vision, spoken language is significantly more typical, as is the perfective aspect and a first person subject. These characteristics signal that Vision is an action that is often mentioned in conversation once, typically said about the self and is something that is typically successfully completed. From a constructional point of view, Vision is more often directed at something (look at something) than just about perceiving something (see something).

(18) Я 25-го еду в Свердловск/ а 27-го в Киев уезжаю/ здорово/ подружку свою увижу/ по магазинам похожу/ у нас щас [sic] одна проблема// [Разговор родственников // Живая речь уральского города, 1990]

I am going to Sverdlovsk on the 25th and the 27th I'm leaving for Kiev/cool/I'll see my friend/I'll go shopping/we only have one problem now//

Touch verb prefers imperfectives more strongly, third person subjects (i.e., it is said about others), negated contexts and non-literal use. In other words, Touch is typically found in ongoing or repeated situations, that are often negated (don't touch!) and in situations where no physical sensors are involved. Construction-wise, touch behaves similarly, with a preference for "feel (as/like) x" statements.

(19) Недели три его никто не трогает, вообще забыли, что есть такой. [Анатолий Трушкин. 208 избранных страниц (1990-2002)]

For three weeks no one has been touching him, they've totally forgotten that there is someone like him.

Hearing favours the imperative (listen!) and is typically used literally (meaning physical receptors are involved). As far as constructional preferences are concerned, hearing occurs significantly more often without direct objects at all, leading to statements such as I am listening that are confirmations of an invisible act.

(20) — Давид, а что будете делать с пришедшими из Чечни боевиками?

David, what do you plan to do with the warriors who've come from Chechnya?

— Слушайте, они ваши, российские, бандиты. Сами с ними воюйте, а нам и собственных хватает. [Елена Лория, Александр Хохлов. Руслан Гелаев может спать спокойно. Корреспонденты «Известий» стали участниками операции грузинских спецслужб в Панкисском ущелье (2002) // «Известия», 2002.09.04]

Listen, they are your, Russian, bandits. Deal with them yourselves, we've got enough of our own.

No significant differences between the 3 groups were found for voice, number or tense. All perception types are predominantly used agentively, all prefer the singular and refer most often to events in the past.

Interestingly, non-literal use is rare for Hearing. This is very different from the situation for Vision and Touch verbs that are often used non-literally. More specifically, Hearing is less likely to be used in situations that do not require agentive sensory perception than Vision is (in fact, Hearing is found used non-literally in 9 instances only), while Touch occurs significantly more frequently in non-literal situations than Vision does (overall 214 times out of 293). There is also a difference in degree of non-literalness between the perception types, with Hearing in (21) being used in a situation where visual perception is needed rather than auditory perception, since the readers will read the opinion, not hear it, but Touch in (22) is instantiated in the emotional domain.

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

(21) Думаю, читателям журнала интересно услышать мнение конструкторов, а лучше диалог по преимуществу их «питомцев». [Морские хищники на сухопутье (2004) // «Солдат удачи», 2004.04.07]

I think it would be interesting for the readers of the journal to hear the opinion of the builders (…)

(22) — Зато на флангах Фигу и Роберто Карлос чувствовали себя вполне свободно… [Юрий Семин: «Ничья с «Реалом» -- это не предел» (2002) // «Известия», 2002.12.13]

- But at the flanks Figo and Roberto Carlos felt completely free …

3.1.2 Agentive versus experiential perception

A similar analysis as for the perception types can be carried out for agentive versus experiential perception. This analysis takes as input the two columns in the top row of Table (1) and compares the behaviour of agentive perception with experiential perception, contrasting data for the individual verbs per perception type.

At this level too, there are striking similarities but also significant differences between the ways in which the three agentive and the three experiential verbs are used. Overall, Hearing seems to be the most neutral perception type with both agentive and experiential Hearing behaving nearly as expected (in the statistical hypothesis testing sense of the word, i.e. as expected if there would not be an association between the variables in question). Vision showed most deviations for the agentive verbs and Touch for the experiential verbs.

For agentive perception verbs it is the behaviour of *smotret’* that diverges significantly in half of the properties tracked. In contrast to what would be expected, *smotret’* prefers the perfective aspect and is more readily used as a gerund than the other verbs, likes the first person and favours non-past forms. In other words, as summarized in Section 3.1.1, *smotret’* expresses an action that is said more often than expected by chance about the self and is predicted to happen once or be accomplished in the future.

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

The differences between *smotret'* and the agentive Hearing and Touch verbs imply that the agentive seeing concept as expressed by *smotret'* in Russian is delineated quite well linguistically and presented as distinct from Hearing and Touch. This makes it easier to pick up how to use *smotret'* from language alone, but it also makes it more difficult to draw parallels with Hearing and Touch and to transfer experiential information from Hearing and Touch to Vision.

As far as experiential perception is concerned, the experiential Vision verb *videt'* diverges from expectation on only one property; in all other cases it is Touch or Touch and Hearing that diverge. Experiential seeing can thus by and large be copied from experiential hearing, but the situation with experiential feeling is different. *Čuvstvovat'* in particular occurs more often than expected by chance in the imperative (and less often in the participle), in the present tense, with a third person subject and in a negated sentence. This is in line with the properties listed as specific for Touch in general in Section 3.1.1.

3.1.3 Agentive versus experiential Vision

This type of comparison can also be applied to the members of each pair of verbs individually. Browsing the data we see that there are hardly any morphological or syntactic properties that are exclusive to one verb or the other. Instead, all verbs display the same range of morphological properties and syntactic possibilities. Yet it remains possible to distinguish between the verbs in each pair by tracking their preferences for use in or with a specific grammatical context.

Comparing the core verbs *videt'* and *smotret'* property for property reveals in which respects the two Vision verbs are similar or differ on a number of morphological and syntactic properties.

Videt' occurs more often than expected in the imperfective aspect; it is directed at a direct object and often used non-literally. It also occurs more often than expected in the passive voice, and as in the indicative and as a participle. Finally, it is something that is more frequently than expected said about the past and about singular subjects.

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

Smotret' encodes an action that is attracted to the perfective aspect; it is directed at an object using a preposition and is typically used literally. *Smotret'* gravitates more than expected by chance towards the imperative and infinitive, modes that do not exist in the past or are not tensed in Russian.

3.2 Multivariate analysis: a tree and forest model

After an initial univariate analysis that has revealed which properties are more strongly attracted than expected by chance to individual verbs as well as to certain groupings, such as agentive/experiential or type of perception, we are now ready to move on to answering the question of whether perception type can be predicted from the morphological and syntactic variables.

A tree and forest model that relies on recursive partitioning is used to determine whether and how well the perception type mentioned in a sentence can be predicted on the basis of the morphological and syntactic properties available. A single classification tree shows how much of each perception type is accounted for (in the order Vision/Hearing/Touch) by tracking a specific variable level; this makes classification trees particularly suited for the purposes of this paper. A tree model discards non-significant predictors automatically and naturally allows for interactions. The resulting visualization is easy to read and provides straightforward insights into the structure of the data. Finally, there is a mechanism available for validating the proposed tree model. The classification forest relies on bootstrap samples, that is, samples of size N drawn with replacement from the original dataset with N observations. Using the R party package both a classification tree and a classification forest were constructed, with the forest grown from 1000 random samples and number of variables to consider at each split set to 4.

A classification tree provides an optimal partitioning of the data and presents a procedure for deciding whether the perception type expressed in a sentence will be, in this case, Vision, Hearing or Touch. The classification tree for the perception data, based on all variables

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

presented in Table (3) with the exception of literal vs. non-literal reading, is represented in Figure (1).

INSERT FIGURE 1 ABOUT HERE (ctree seed 11.pdf)

Figure 1. Classification tree for perception data

Each split in the tree is labelled with a decision rule. Before determining the rule at each node, the algorithm inspects all predictors and selects the one that is most useful. The algorithm does not look ahead, however, and cannot consider decisions that would yield a slightly worse split locally but would do significantly better globally. The first split (in the oval) in this tree is on Construction type. The accompanying p-value indicates that Perception types are well separable if the construction type is known. In this particular case, the partitioning shows that if the constructions consist of a target verb followed by a noun, adjective or adverb, then follow the right branch. This branch leads directly to a terminal node that unites 51 occurrences of Touch, at the bottom right hand side; these can be predicted directly from their occurrence in these 3 constructional circumstances. The bar graph shows the total number of tokens in the node and how they are divided over the 3 Perception types. Each leaf node contains a unique subset of the data and the leaf nodes jointly make up the entire dataset.

Bearing in mind how classification trees work, this part of the data has now been cordoned off and further decision rules will focus on splitting up the remainder of the data as well as possible. The search for the locally best performing splitting criterion is now repeated for the remainder of the data. At each next branch a new decision rule is presented that directs us further down the tree past always purer nodes; the realization of Vision vs. Hearing vs. Touch should be more pure or extreme in the daughter nodes than in the mother nodes higher up. In this diagram, Construction type shows up as a significant predictor of Perception type at the second split as well. If the target verb is used as adverb or is followed by an adverbial or prepositional clause, it is very likely to be a Vision verb, as shown in the terminal node in the left-hand corner at the bottom of the diagram that contains 112 instances.

The tree in Figure (1) reveals that constructional properties play a crucial role in distinguishing between the perception types: they are found highest up in the tree, executing the first and second splits. The property that comes into play next for splitting the remainder of the data as best as possible is polarity; depending on the polarity of a sentence, either verb number or spoken vs written situation is required to achieve further splitting. If the sentence is negated, only verb number is needed to arrive at an endnode, with plural verbs in negated contexts often being instances of Touch.For singular verbs the preference is less pronounced. The distinction between spoken vs written language plays an important role in positive contexts: Given that nearly half of all instances of Hearing are grouped under the right branch that concerns written language, we can conclude that Hearing is a type of event that needs reporting on in written language in particular. Yet, the fact that part of the Hearing data is cordoned off by its non-occurrence in spoken language does not mean that Hearing would be the situation least frequently talked about. In fact, Touch is least frequently used in spoken contexts, but half of Touch has already been accounted for in earlier nodes, hence Touch is less visible in the remaining leaf nodes. Overall, Vision is used three times more often in speech than Touch is.

All in all, the tree correctly classifies 56.9% of all instances, with 207/300 Hearing, 162/300 Vision and 139/293 Touch correctly classified (Table (4)). This is twice as good as randomly choosing, which would yield 1/3 correct. Moreover, the perception types are most often predicted as themselves, i.e. the highest values are on the diagonal in the table, signalling that the classification accuracy is good for all three perception types involved.

Table 4. Classification according to the standard classification tree

	Vision	**Hearing**	**Touch**	**Total**
Vision	**162**	70	24	**256**

Hearing	121	**207**	130	**458**
Touch	17	23	**139**	**179**
Grand Total	**300**	**300**	**293**	**893**

That being said, Touch is less well predicted than the two other senses and is least frequently predicted overall. But, in this case of how grammar relates to meaning incorrect classifications are also revealing. Instances of Vision and Touch that are mispredicted are most often predicted as Hearing, while mispredictions of Hearing are more or less equally divided between Vision and Touch. A prediction mechanism working on the basis of morphological and syntactic properties alone is unable to distinguish well between Vision and Hearing. Hearing is also the most frequently predicted Perception type, stressing that it occupies the most neutral position when viewed in terms of morphological and syntactic properties alone.

A single tree is likely to overfit the data, however; growing a forest based on resampling mitigates against this risk. A random forest also makes more precise predictions than a standard classification tree. In this particular example, a random forest of 1000 trees increases the correct prediction rate to 64.5% with 193/300 Hearing, 172/300 Vision and 211/293 Touch correctly classified (Table 5).

Table 5. Classification according to the random forest

	Vision	**Hearing**	**Touch**	**Total**
Vision	**172**	50	20	242
Hearing	81	**193**	62	336
Touch	47	57	**211**	315
Grand Total	**300**	**300**	**293**	893

Generally, each of the three perception types is most frequently predicted as itself. Hearing remains the most overpredicted perception type, but it is now Vision that is least frequently predicted. If Vision and Touch are mispredicted, they are most frequently predicted as Hearing, while Hearing is almost equally divided between Vision and Touch.

On the basis of the forest, the importance of each variable can be calculated. Using random permutation of the labels of each variable, the relative importance of the different predictors for the classification accuracy of the model is assessed. The result is shown in Figure (2).

INSERT FIGURE 4 ABOUT HERE (dotchart seed 11111 runs 1000.png)

Figure 2. Variable importance plot

The variable importance plot shows that construction is the strongest predictor (at 0.059, rounded up), followed by negative or positive polarity (at 0.032), spoken or written situation (at 0.020) and aspect (at 0.014). Leaving the remaining variables out reduces the prediction accuracy of the model with less than 1%. Verb number, form, person, tense and voice are thus rather fine-tuning variables than anything else, and each taken individually, they do not contribute much to a correct classification of situations as types of Perception.

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

4. Discussion

This paper set out to determine to what extent morphological and syntactic criteria could aid learners in setting up concepts of Perception. How do the findings described in Section 3 relate to the hypothesis put forward in the literature that learners are be able to infer much of the meaning of a verb from the set of constructions it is used in.

Could speakers rely on differences in use to distinguish the 3 main types of perception from each other?

Yes. Data from the written and spoken subsections of the RNC shows that the 3 perception types are characterized by a different statistical distribution of morphological and syntactic forms in usage.

In order for a speaker to distinguish Vision from Hearing and Touch on the basis of grammatical information alone, s/he first and foremost needs to track the range of constructions in which the verbs occurs. The relation between argument structure constructions and a verb's meaning is well documented (cf. Apresjan 1967 for Russian, Levin 1993 for English) and it comes as no surprise that differences in the types of constructions a verb occurs in would be a strong predictor for a verb's meaning. Judging from these properties alone, 64.5% is correctly predicted, with 193/300 Hearing, 172/300 Vision and 211/293 Touch correctly classified. Generally, Hearing is the most strongly overpredicted on this subset of criteria, and Vision is least frequently predicted.

Could speakers rely on similarities in use to be able to map Hearing and Touch experiences to Vision?

Yes, the data show that this is indeed the case, in particular for experiential perception. *Videt'* is the verb that is used in a way that does not deviate significantly from what is expected within the confines of the set of basic perception verbs analysed here. It appears to be very much like Hearing, at least from a linguistic point of view.

Agentive perception is very different, with look diverging significantly from touch and listen as far as morphology and syntax are concerned. Language is rather specific about the difference between agentive Vision and Hearing/Touch; in contrast to what would be expected. S*motret'* prefers the perfective aspect, occurs more often than the other verbs as gerund, attracts the first person and favours non-past forms. In other words, as summarized in Section 3.1.1, *smotret'* expresses an action that is typically said about the self and is predicted to happen once or be accomplished in the future; this is not so for Hearing/Touch. In addition to providing an estimate of the contribution language makes to concept formation, this dataset contains the necessary information to complement Landau & Gleitman's (1985) study.

On the basis of a detailed analysis of the way in which a blind girl, Kelli, acquired and used verbs of seeing, Landau & Gleitman (1985) concluded that the precise meaning of a specific verb could be determined on the basis of the set of morphological and syntactic contexts it is used in. The fact that "a large number of semantic distinctions is encoded onto a limited number of phrase organizations of the clause" results in a "many-to-few mapping of the meaning components onto the surface forms". In other words, there is very little information in any single syntactic format that is attested for a verb, because that format serves many distinct uses. This implies that "the child can make no sure induction from a single phrase structure to the meaning it encodes and vice versa." However, the set of subcategorization frames associated with a verb is highly informative about the meaning it conveys. As Landau and Gleitman put it, each verb entry includes a set of subcategorization frames that do double-duty and also represent part of the semantics of each verb. This finding is in line with the distributional hypothesis that was originally proposed by Harris (1954). Harris' distributionalism predicts that the meaning of a word is derived from the linguistic contexts in which it occurs, with words that are more similar occurring in a larger number of shared contexts.

Although Landau & Gleitman (1985: 138-142) promoted the idea that linguistic experience, and in particular the sets or networks of constructions a verb is used in, may be a crucial contributing factor for blind children to get a grip on verbs such as look and see, on their account, sensory experiences from Hearing and Touch would need to be transferred to Vision to account for others' knowledge: experiential analogies with cognate sensory experiences such as hearing and touching would be needed to fully reconstruct seeing in the blind (Landau & Gleitman 1985: 83). Although Landau & Gleitman point out that what is at stake are experiences in the hearing (listen vs hear) and touching (touch vs feel) domains that have clear parallels in the vision domain (look vs see), they remain silent about how these parallels would be identified and the transfer would happen.

With the data presented in this paper, we can explore the idea that the proposed transfer of experience from Touch and Hearing to Vision would be facilitated if guided by a shared linguistic structure. For example, experience of the agentive-experiential contrast in the domains of hearing (listen vs hear) and touching (touch vs feel) could shape knowledge of the agentive-experiential divide (look vs see) in the vision domain. This distinction could be picked up by registering which verbs occur in the imperative (look, listen, touch) and which ones do not (see, hear, feel), across all three domains, and as such provide more precise guidance on what needs to be mapped across domains.

Before setting out, a caveat needs to be expressed. In a sense, this part of the discussion represents a thought experiment; it is not about language as used by or with blind speakers, but about a snapshot of language in general, about what language offersthat (blind) speakers could make use of. Yet, the lack of visual input in the blind could change the distribution of the inputs from the other sensory channels (known as sensory compensation), and this could possibly result in behavioral differences between the sighted and the blind.

Could speakers rely on similarities in use to be able to ground Vision indirectly in perception and hence to some extent obviate the need for sensory experience in concept formation?

Given a large enough shared linguistic basis, analogies drawing on comparable sensory experiences involving cognate sensory domains could be directly accommodated. Especially the experiential perception verbs lend themselves well to supporting transfer from a cognate sense to Vision and enabling blind speakers to infer that Seeing is another sense, just like Hearing/Feeling. Overall, the picture any speaker gets from linguistic data alone is that Seeing is something very similar to Hearing and Feeling, but that the act of Looking differs in many respects significantly from Listening and Touching.

Non-literal use of perception verbs plays a potentially important role here. From experience with Touch verbs, speakers are accustomed to perception verbs being used in a non-literal way, when sensors are not activated. Non-literal use could form a pathway for blind speakers to build up a more general concept of Vision, without perceptual experience, as it decouples physical sensations from the verbs for which the blind have perceptual experiences and creates pathways for supporting cross-modal analogy.

Could speakers rely on differences in use to be able to use Vision verbs correctly?

Yes. The Vision verbs are characterized by distinct typical usage patterns, both as far as morphology and syntax are concerned. To recapitulate:

Videt' occurs more often than expected in the imperfective aspect; it is said about past events and about singular subjects. In other words, *videt'* encodes an action that goes on for some time and is typically commented on after it has taken place. *Videt'* also occurs more often than expected in the passive voice, used as in the indicative and as a participle; it is an activity that someone carries out and is presented as a fact (indicative) or as a property (participle) of the subject. *Videt'* is directed at a direct object but is often used non-literally, indicating visual sensors need not be activated when the verb is used.

Smotret', its agentive counterpart, encodes an action that is attracted to the perfective aspect and gravitates more than expected towards the imperative and infinitive, modes that do not exist in the past or are not tensed at all in Russian. *Smotret'* thus seems to be used, more often than expected, in situations where someone else is told to perform an act of looking. Unlike *videt'*, *smotret'* is directed at an object using a preposition and is typically used literally.

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

5. Conclusion

This corpus study yields a first estimate of the potential that a speaker's experience with language has for contributing to sensory concept formation in Russian. In more than 64% of all usage instances of perception verbs, the perception type referred to can be classified using a model that relies on morphological and structural information alone.

Although overall verbs expressing Vision, Touch and Hearing occur with the same types of morphological markers and participate in by and large the same sets of syntactic constructions, the frequency with which they occur in each one is different, leaving a distinct path of preference. This requires a refinement of the bootstrapping hypothesis, turning it more distributional than was originally envisaged: Speakers do not only need track the sets of constructions a verb occurs in, but also how frequently a particular verb is encountered in a specific construction or with a particular morphological marker (cf. Newman 2008; Divjak 2006, 2010; Janda & Lyashevskaya 2011).

There is further evidence that in Russian, like in English, Vision verbs can be distinguished from one another on this same principle, i.e. on the basis of the frequencies with which they are encountered with certain morphological properties and in certain syntactic constructions.

From this analysis, it can also be concluded that in Russian, Hearing and Touch together could help scaffolding concepts of Vision in the absence of vision. There are strong parallels between the structures in which Vision and Hearing/Touch participate, which could support the inference that Vision is "another sense like Hearing/Touch". At the same time, Touch is frequently used non-literally, and the verb is not restricted to expressing Touch (occasionally also being used to refer to smell and taste), which paves the way for decoupling physical sensations from the verbs and creates pathways for supporting cross-modal analogy.

It is expected that an analysis that takes into account the lexical semantics of the elements in each sentence, i.e. a full Behavioral Profile study of the verbs involved (Divjak 2010 and references therein), would further improve the classification accuracy and work on this is under way. **Notes**

The final publication is available at Benjamins via https://benjamins.com/catalog/fol.22.1.03div

* I would like to thank the three anonymous FoL reviewers as well as Maria Ovsjannikova for valuable observations and suggestions.

[1] Smell and taste have traditionally been considered inferior, that is have been said to play a less prominent role in humans, but see Wnuk & Mahid (2012) for a recent reassessment.

[2] An alternative that is less frequently encountered in the corpus is *potrogat'* ('touch'), which occurs 714 times in the written subcorpus and only 30 times in the oral corpus (last checked on 07.02.2013). The lower frequency of occurrence for the perfective variant *potrogat'* is possibly due to it being more limited to expressing physical sensation than its counterpart *tronut'*, as native speakers report.

[3] A reviewer insisted that osjazat' would be a more appropriate experiential candidate for Touch since it relates to osjazanie, a psychological term for tactile experience. The latter being true, the verb *osjazat'* is not only highly specific (comparable to referring to vision as ophtalmoception) but also not very frequent, in particular in spoken language (there are about 300 occurrences in the written 130 million word corpus and 8 in the spoken 10 million word corpus). This makes it unlikely for average non-highly educated speakers to know it.

[4] In fact, it is also possible to hear a smell in Russian.

[5] Four of these used *čuvstvovat'/počuvstvovat'* to encode smell, while three were homomyms, with tronu representing a form of the noun 'throne' rather than of the verb 'touch'.

[6] I am using the terms participle here to refer to the Russian *pričastie* (adjectival participle) and the term gerund to refer to *deepričastie* (adverbial participle). This is standard practice (compare Janda & Lyashevskaya 2011) and does not imply that the categories would behave identically in English and in Russian.

[7] The level "none" used with tense, number and person signals that this variable does not apply to a certain form, e.g. infinitives in Russian are not marked for tense, number and person and would be tagged "none" for all three.

[8] Janda & Solovyev (2009) report that, based on their data, usually only 6-10 constructions are needed to accurately represent the constructional profile of a verb.

[9] As one reviewer pointed out, some of the perception verbs have developed non-perceptual meanings, such as look being used to mean understand (as in *Я смотрю*, *ты уходишь* 'As far as I understand, you are leaving') or wait (*Посмотрим*, *что он скажет* 'Let's wait for his opinion') in certain contexts, and some have grammaticalized (*Смотря что выбирать* 'Depends on what to choose'.) There were no examples of these constructions in the corpus sample analysed, and as such, they will remain outside the scope of the discussion.

References

Apresjan, Jurij D. 1967. *Экспериментальное исследование семантики русского глагола* [Eksperimental’noe issledovanie semantiki russkogo glagola]. Moskva: Nauka.

Divjak, Dagmar. 2003. On trying in Russian: a tentative network model for near(er) synonyms. “Belgian Contributions to the 13th International Congress of Slavicists, Ljubljana, 15-21 August 2003”. Slavica Gandensia 30. 25-58.

Divjak, Dagmar. 2006. Ways of Intending: Delineating and Structuring Near-Synonyms. In Stefan Th. Gries & Anatol Stefanowitsch (eds.), Corpora in cognitive linguistics. Corpus-based Approaches to Syntax and Lexis, 19-56. Berlin: de Gruyter. [Trends in Linguistics].

Divjak, Dagmar. 2010. Structuring the Lexicon: a Clustered Model for Near-Synonymy. Berlin: de Gruyter.

Divjak, Dagmar & Stefan Th. Gries. 2006. Ways of Trying in Russian. Clustering Behavioral Profiles. Journal of Corpus Linguistics and Linguistic Theory 2 (1). 23-60.

Harris, Zellig. 1954. Distributional structure. Word 10 (23). 146–162.

Hofmann, Thomas. 1999. Probabilistic Latent Semantic Analysis. Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI'99).

Karlsson, Fred. 1986. Frequency Considerations in Morphology. Zeitschrift fur Phonetik, Sprachwissenschaft und Kommunikationsforschung 39. 19-28.

Landau, Barbara & Lila Gleitman. 1985. Language and experience. Evidence from the blind child. Cambridge, MA: Harvard University Press.

Landauer, Thomas K., Peter W. Foltz & Darrell Laham. 1998. Introduction to Latent Semantic Analysis. Discourse Processes 25. 259-284.

Levin, Beth. 1993. English verb classes and alternations: a preliminary investigation. Chicago: University of Chicago Press.

Li, Jianguo & Chris Brew. 2008. Which are the best features for automatic verb classification? In Proceedings of ACL-08: HLT, Columbus, Ohio, USA, 434–442.

Louwerse, Max M. 2011. Symbol interdependency in symbolic and embodied cognition. TopiCS in Cognitive Science 3. 273-302.

Louwerse, Max M. & Patrick Jeuniaux. 2010. The linguistic and embodied nature of conceptual processing. Cognition 114. 96-104.

Matthews, P.H. 2007. The concise Oxford dictionary of linguistics (2nd edition). Oxford: OUP.

Nesset, Tore, Laura A. Janda, Julia Kuznetsova, Ol'ga Lyashevskaya, Anastasia Makarova Svetlana Sokolova. 2008. Why poslushat', but uslyshat'? Poljarnyj Vestnik 11. 38-46.

Newman, John. 2008. Aiming low in linguistics: Low-level generalizations in corpus-based research [downloaded from http://www.johnnewm.org/downloads/; page last accessed on 07.06.2013]

The Oxford Russian-English Dictionary; 2nd Revised edition. 1993. Paul Falla, Marcus Wheeler & Boris Unbegaun (eds.). Oxford: Oxford University Press.

Paducheva, Elena. 2004. *Динамические модели в семантике лексики.* Moskva: Языки славянской культуры.

Paradis Carita & Mats Eeg-Olofsson. 2013. Describing Sensory Experience: The Genre of Wine Reviews. Metaphor & Symbol 28 (1). 22-40.

RNC. *Национальный корпус русского языка*. [http://ruscorpora.ru/en/index.html – data extracted July 2012].

Riordan, Brian & Michael N. Jones. 2011. Redundancy in Perceptual and Linguistic Experience: Comparing Feature-Based and Distributional Models of Semantic Representation. TopiCS in Cognitive Science (3). 303–345.

Sahlgren, Magnus. 2008. The Distributional Hypothesis. Rivista di Linguistica 20 (1). 33-53.

Šteinfeldt, Evi. 1970. Russian word count. Moscow: Progress.

Vigliocco, Gabriella, Lotte Meteyard, Mark Andrews & Stavroula Kousta. 2009. Toward a Theory of Semantic Representation. Language and Cognition 1(2). 219–247.

Wilson, Margaret. 2002. Six views of embodied cognition. Psychonomic Bulletin and Review, 9 (4). 625-636.

Wnuk, Ewelina & Asifa Majid. 2012. Olfaction in a hunter-gatherer society: Insights from language and culture. In Naomi Miyake, David Peebles & Richard P. Cooper (eds.), CogSci 2012: Proceedings of the 34th Annual Conference of the Cognitive Science Society (pp. 1155-1160). Austin, TX: Cognitive Science Society.

[9002 words including footnotes and references]