Research Article

Přemysl Mácha*, Terhi Ainiala, Urszula Bijak, Alina Bughesiu, Ivana Crljenko, Oliviu Felecan, Matjaž Geršič, Veronika Gondeková, Michaela Koščová, Tiina Laansalu, Ján Mačutek, Sanda Rapa, Katalin Reszegi, Pavel Skorupa, Ene Vainik, Maya Vlahova-Angelova, Maria Vrachionidou and Halyna Zymovets

Different languages, different landscapes? Exploring linguistic differences in landscape conceptualizations

https://doi.org/10.1515/opli-2025-0081
Received October 17, 2025; accepted January 13, 2026; published online March 2, 2026

Abstract: The goal of the research was to assess whether landscape conceptualization systematically differed across European languages in the context of the linguistic relativity debate. We examined the generalizability of previous findings and evaluated the limits of used methodologies, drawing on a large and diverse sample of 14 languages across 5 linguistic families in Eastern Europe (Slavic, Baltic, Uralic, Romance, Hellenic). Participants were asked to free-list terms associated with three linguistic domains: landscape, animals, and body parts. Mixed ANOVA, pairwise comparisons, frequency distribution, cognitive salience, and semantic network analyses were used to test for relationships and to visualize patterns and differences between languages. We confirmed that the landscape domain was less structured than the domains of animals and body parts. Furthermore, the landscape domain had weaker semantic connections across languages. However, we did not find any systematic differences in landscape conceptualizations across languages that could be clearly attributed to linguistic factors. Rather, we argue that the observed variability is more likely the result of multiple factors – geographical, cultural, and linguistic. Building on previous research, we propose more nuanced methods in future research that integrate qualitative, ethnographic insights with quantitative methods, while accounting for extra-linguistic factors.

***Corresponding author: Přemysl Mácha**, Institute of Ethnology, Czech Academy of Sciences, Brno, Czech Republic, E-mail: macha@eu.cas.cz. https://orcid.org/0000-0001-9985-1064
Terhi Ainiala, University of Helsinki, Helsinki, Finland. https://orcid.org/0000-0003-4998-9983
Urszula Bijak, Institute of Polish Language, Polish Academy of Sciences, Kraków, Poland. https://orcid.org/0000-0003-0378-6929
Alina Bughesiu and Oliviu Felecan, Universitatea Tehnica Cluj-Napoca, Cluj-Napoca, Romania. https://orcid.org/0000-0001-8403-4432 (A. Bughesiu)
Ivana Crljenko, University of Zadar, Zadar, Croatia
Matjaž Geršič, ZRC SAZU, Anton Melik Geographical Institute, Ljubljana, Slovenia. https://orcid.org/0000-0001-9640-6037
Veronika Gondeková, Matej Bel University, Banská Bystrica, Slovakia
Michaela Koščová, Mathematical Institute, Slovak Academy of Sciences, Bratislava, Slovakia. https://orcid.org/0000-0002-5541-1224
Tiina Laansalu and Ene Vainik, Institute of the Estonian Language, Tallin, Estonia. https://orcid.org/0000-0002-4320-8755 (E. Vainik)
Ján Mačutek, Mathematical Institute, Slovak Academy of Sciences, Bratislava, Slovakia; and Department of Mathematics, Constantine the Philosopher University in Nitra, Nitra, Slovakia. https://orcid.org/0000-0003-1712-4395
Sanda Rapa, University of Latvia, Riga, Latvia. https://orcid.org/0000-0002-7959-9719
Katalin Reszegi, University of Debrecen, Debrecen, Hungary. https://orcid.org/0000-0002-4443-3284
Pavel Skorupa, Institute of the Lithuanian Language, Vilnius, Lithuania. https://orcid.org/0000-0003-3903-7006
Maya Vlahova-Angelova, Institute for Bulgarian Language, Bulgarian Academy of Sciences, Sofia, Bulgaria. https://orcid.org/0000-0002-7826-1898
Maria Vrachionidou, Research Center for Modern Greek Dialects, Academy of Athens, Athens, Greece. https://orcid.org/0009-0008-9168-9216
Halyna Zymovets, O. O. Potebnia Institute of Linguistics, National Academy of Sciences, Kyiv, Ukraine. https://orcid.org/0000-0002-6232-7469

Keywords: landscape perception; linguistic relativity; Eastern Europe; semantic differences

1 Introduction

Understanding how people perceive and conceptualize landscapes has become an increasingly important topic for fundamental and applied research. We are facing the need to design adequate landscape policies in a world dramatically and rapidly transformed by climate change, migration, urbanization, desertification, deforestation, glacier melting, and the appearance of extensive solar and wind farms. Various factors shape people's perception of landscape, one of them being language. In a groundbreaking study, van Putten et al. (2020) tested the influence of language on people's conceptualization of landscape on a sample of seven Germanic and Romance languages spoken in Western and Northern Europe. The researchers found that conceptualizations of landscape differed across languages and between language speakers, concluding that local conceptualizations of landscape should be taken into account when designing national or European landscape policies.

In addition to clear practical applications, these findings can also be placed in the context of the ongoing debate on linguistic relativity, i.e. the influence of language on cognition (Malt and Majid 2013). This debate has a long and rich history and is far from finished (Danesi 2021; Levinson 1996; Lucy 2016). Numerous studies have shown that conceptualizations of the physical environment vary greatly across languages and there is a large cognitive and linguistic diversity in landscape categorization and spatial grammar (Johnson and Hunn 2010; Levinson 2003; Levinson and Wilkins 2006; Mark et al. 2011). However, no conclusive evidence has been found either in natural or experimental research design settings which would prove or disprove the effect of language on spatial perception (Hansen and O'Meara 2020; Heegård and Liljegren 2018; Urban 2020). Still, there appears to be an emerging consensus that language may perhaps not determine thought and spatial cognition ("strong relativity hypothesis") but it channels thought into certain schemas, promotes specific forms of spatial perception, orientation and conceptualization, and makes real-life spatial tasks easier to perform ("weak relativity hypothesis") (Baier et al. 2023; Haun et al. 2011; Wolff and Holmes 2011).

The purpose of this article is to contribute to the linguistic relativity debate by building on van Putten et al.'s study. We draw on a sample of 14 languages from 5 language families (Slavic, Baltic, Uralic, Romance, Hellenic) spoken in the eastern half of Europe from Finland in the north to Greece in the south. After presenting our findings and comparing them with van Putten et al.'s study, we critically evaluate the limits of our methodology and suggest its adaptation to better serve comparable research in the future.

2 Methods

2.1 Research sample

We collected usable data from a total of 2,970 speakers of Bulgarian, Croatian, Czech, Estonian, Finnish, Greek, Hungarian, Latvian, Lithuanian, Polish, Romanian, Slovak, Slovene, and Ukrainian living in the respective countries. Research participants were contacted via personal and professional networks and social media. The most important criterion for selection was that the participants had to be native speakers of the given language. However, we tried to diversify the sample to approximate the structure of the general population in terms of residence, age, and educational level to account for potential sociodemographic differences in lexical proficiency and semantic use. While the age structure of our sample was relatively balanced, women and university-educated respondents were heavily overrepresented compared to the general population. The residential structure of our sample varied across countries and included respondents living in villages, small towns, and large cities (see Table 1 for sample composition).

Table 1: Sample composition.

Country	*N*	Age (mean)	Age (range)	% Women	% With a university education	% Villagers
Bulgaria	130	43	19–84	66	85	8
Croatia	243	40	13–72	73	78	20
Czechia	160	34	17–72	73	67	29
Estonia	247	49	19–83	87	84	18
Finland	319	46	14–80	88	82	19
Greece	115	41	12–67	83	77	7
Hungary	209	30	12–76	76	43	20
Latvia	504	41	12–72	90	84	22
Lithuania	123	40	11–73	85	72	6
Poland	175	34	17–80	67	63	22
Romania	151	31	17–75	84	58	28
Slovakia	134	33	16–68	75	65	32
Slovenia	226	45	15–81	73	91	47
Ukraine	234	34	13–71	86	78	15

2.2 Data collection

Participants were asked to fill out an online form modelled after van Putten et al.'s study (van Putten et al. 2020) to allow for comparison. The collection of responses took place between March 10 and April 4 2023. The need for informed consent was waived by the ethics committee of the Institute of Ethnology of the Czech Academy of Sciences because the questionnaire was fully anonymous, no personal or sensitive information which would make it possible to identify specific participants was collected, and the questions themselves did not present any potential emotional or psychological harm to participants. The research complied with all the relevant national regulations, institutional policies and in accordance with the tenets of the Helsinki Declaration, and was approved by the authors' institutional review board or equivalent committee.

The form consisted of four sections. In the first section, participants were asked to free list words which they associated with three linguistic domains – landscape, body parts, and animals (see Table 2 for translation of these domain names into the research languages). The participants were given 3 min to list as many terms as possible for each domain. In the second section, the participants were asked, using a five-point scale, to rate how difficult it was to list terms in each of the three domains. In the third section, we asked how many hours per week the participant spent engaging in outdoor activities (none/1–2 h/3–5 h/6–8 h/9 and more) and how often the participant visited various landscape objects (high mountains, low mountains, forest, countryside, river, lake, coast, desert, volcano, island, city park) choosing from predefined options (every day/at least once a week/at least once a month/at least once a year/at least once every 5 years/have been at least once/never visited). For the same list of landscape objects the participants were then asked to indicate which of them they had visited in the past six months. In the last section, we asked about the sociodemographic profile of the participant – age, gender (male/female/other), residence (village, small town, big city) and highest level of educational attainment (elementary/secondary/university). This section served as a reference to compare our sample to the general population but also to test for potential differences between sociodemographic groups irrespective of their language.

The form was created in English and then it was translated into individual languages by the members of the research team. This process was an extremely challenging task because some landscape categories including the landscape domain itself proved very difficult to translate in a way that would ensure identical understanding not only within each country but across the entire sample. For example, *krajina*, the immediate Slovak equivalent of the English term *landscape*, also means *country* (e.g. Slovakia). To avoid confusion, the term *krajinná scenéria* (landscape scenery) was used instead. In Croatian, three different terms for landscape – *krajolik/krajobraz/pejzaž* – are used synonymously. Two different terms had to be used in Ukrainian – *краєвид/ландшафт* – and Bulgarian – *ландшафт/пейзаж* to account for semantic differences between them. In Romanian, the word *peisaj*

Table 2: Translation of domain terms.

Language	Landscape	Body parts	Animals
Bulgarian	ландшафт/пейзаж	части на тялото	животни
Croatian	krajolik/krajobraz/pejzaž	dio tijela	životinja
Czech	krajina	části těla	zvířata
Estonian	maastik	kehaosa	loom
Finnish	maisema	ruumiinosat	eläimet
Greek	τοπίο	μέρη του σώματος	ζώα
Hungarian	táj	testrészek	állatok
Latvian	ainava	ķermeņa daļa	dzīvnieks
Lithuanian	kraštovaizdis	kūno dalys	gyvūnai
Polish	krajobraz	części ciała	zwierzęta
Romanian	un loc natural/o formă de relief	părțile corpului	animale
Slovak	krajinná scenéria	časti tela	zvieratá
Slovenian	pokrajina	del telesa	žival
Ukrainian	краєвид/ландшафт	частини тіла	тварини

which is cognate with the Spanish *paisaje* and French *paysage* was not used because of the polysemy of the word which would have led to an ambiguous understanding of the questions about landscape. Instead, *un loc natural/o formă de relief* was chosen to avoid misinterpretation. Finally, it should be noted that three Slavic languages (Bulgarian, Croatian, and Ukrainian) use terms for landscape derived from Romance (*pejzaž, пейзаж*) and Germanic (*ландшафт*) languages, making the root words' etymology opaque and semantically irrelevant. The terminological situation was thus somewhat more complicated than in van Putten et al.'s study, hinting, at the same time, at the significance of language in shaping the landscape perceptual domain. Analogical challenges were encountered with terms for individual landscape objects (see Discussion).

Before data collection began, the form was tested on a small sample of 10–20 respondents in each country to ensure the correct interpretation of individual questions and to eliminate any remaining ambiguities. After data collection ended, the free-listed responses were standardized. Spelling errors and typos were corrected and upper and lower cases, singular and plural forms, spelling variants, and definite and indefinite articles were unified to make the following statistical analysis possible.

3 Results

3.1 Listing difficulty

Overall, participants considered listing terms in the landscape domain as the most difficult task. Mixed ANOVA was used to test for differences between domains, languages and their combinations. The mean score of difficulty (with 1 as easiest and 5 as most difficult) was $M = 2.0$ for landscape as compared to $M = 1.8$ for body parts and $M = 1.7$ for animals. Although these differences were relatively small, the effect of domain proved to be statistically significant overall, $F(1.94, 5{,}725.32) = 60.875$, $p < 0.001$, as well as in contrasts (landscape × animals $p < 0.001$, landscape × body parts $p < 0.001$ and body parts × animals $p < 0.001$).

However, the differences in perceived listing difficulty were not consistent across all languages. Participants considered listing landscape terms as the most difficult in 11 languages, whereas the speakers of Croatian, Latvian and Polish classified listing these terms as equally or less difficult than listing terms in the other two domains. More importantly, pairwise comparisons showed that the reported difficulty of listing terms in the landscape domain was significantly higher than in the body parts domain only in Czech ($p = 0.010$), Estonian ($p = 0.032$), Lithuanian ($p = 0.035$), Slovak ($p < 0.001$) and Slovene ($p < 0.001$) and higher than in the animals domain only in Estonian ($p < 0.001$), Finnish ($p < 0.001$), Hungarian ($p < 0.001$), Lithuanian ($p = 0.018$), Romanian ($p = 0.041$), Slovak

($p < 0.001$) and Slovene ($p < 0.001$). A detailed analysis thus showed a more complex picture than the summary statistic.

The effect of language was also significant, $F(13, 2{,}956) = 11.094$, $p < 0.001$, and so was the combined effect of language and domain, $F(25.18, 5{,}725.32) = 6.336$, $p < 0.001$. However, since we had no specific predictions on differences in perceived difficulty between languages, we do not comment on this any further.

3.2 Number of terms listed

Participants listed, on average, 7.9 terms for landscape, 8.4 terms for body parts and 10 terms for animals. Mixed ANOVA was used to analyze the observed variation. The effect of the domain for the number of listed terms was statistically significant, $F(1.88, 5{,}560.44) = 170.302$, $p < 0.001$. The participants listed on average the smallest number of terms for the landscape domain in all languages with one exception – in Ukrainian, the average number of responses was lower for body parts than for landscape (see Figure 1). The effect of language was also statistically significant, $F(13, 2{,}956) = 80.347$, $p < 0.001$, as was the combined effect of domain and language (ANOVA, $F(24.45, 5{,}560.44) = 10.470$, $p < 0.001$). As with listing difficulty, we had no specific predictions on differences between languages in the number of listed terms, so we do not comment on it further.

Pairwise comparisons did not show consistent results. In seven languages, the landscape terms were significantly less numerous than the terms for animals (Bulgarian $p = 0.022$, Estonian $p < 0.001$, Finnish $p < 0.001$, Hungarian $p = 0.003$, Lithuanian $p = 0.045$, Romanian $p < 0.001$, and Slovak $p = 0.042$). The difference between the number of terms listed for body parts and animals was significant in five cases (Bulgarian $p = 0.039$, Estonian $p < 0.001$, Finnish $p < 0.001$, Hungarian $p = 0.033$, Romanian $p = 0.010$). The difference between the number of landscape terms and terms for body parts was not statistically significant in any language.

3.3 Frequency distributions

Terms in each domain were listed with different frequencies. Some were unique, others appeared in most responses and many occurred in characteristic pairs, triads or even more complex associations (see Section 3.5 on semantic networks). The analysis of the frequency of individual terms and the co-occurrence of their pairs

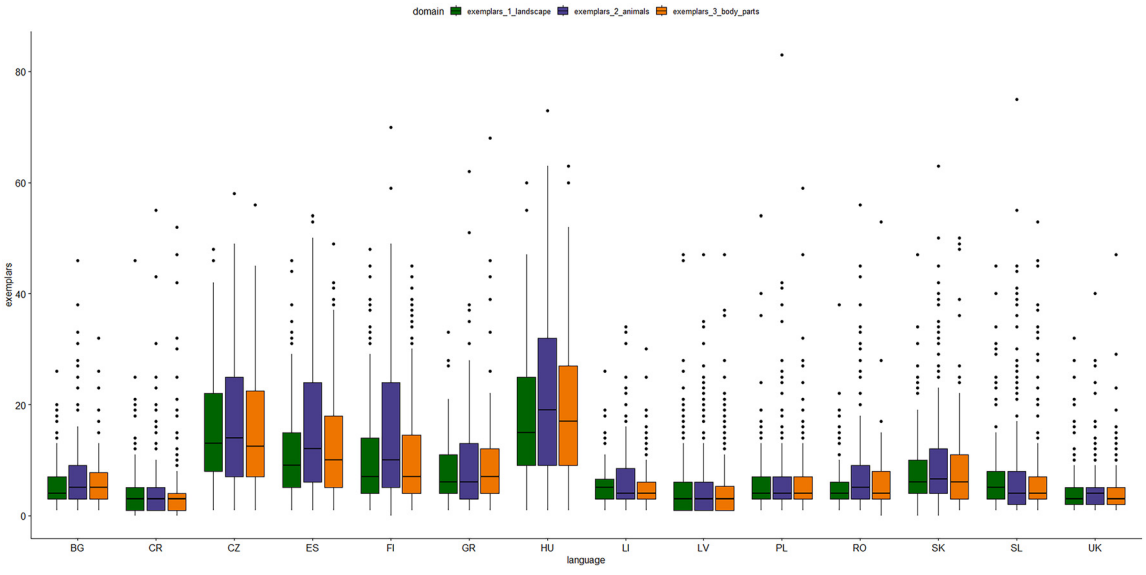

Figure 1: Number of terms listed for landscape, animals and body parts.

(i.e. terms listed adjacently) showed a relatively consistent picture as demonstrated in Figures 2 and 3. Only a few terms across all domains and languages were shared by most participants and their number drops rapidly from left to right as terms with lower frequency and unique terms begin to dominate. Except for Romanian, landscape terms shared among participants occurred with the least frequency compared to terms associated with the body parts and animals domains. This trend was even more pronounced for the frequency of co-occurrence pairs, in this case including Romanian. In other words, there was the least consensus among participants on what constitutes the landscape domain. However, the tails of landscape curves were not longer than the tails of body parts and animals curves, indicating a high incidence of low frequency and unique terms across the three domains. Also, the landscape curves were not consistently steeper or less steep than the curves for the other two domains – this property varied by language.

3.4 Cognitive salience

A cognitive salience score was calculated for each standardized term listed in the three domains. This measure is used to analyze data generated in free-listing exercises. It reflects the extent to which a free-listed term is shared (its frequency) and what priority it is given (position in the list). The formula used to calculate cognitive salience scores was S = F/(NR) where F is frequency, *N* the number of participants, and R the mean rank (Sutrop 2001; van Putten et al. 2020).

Tables 3–5 show the five most salient terms for each domain and language, including their translation into English. The visualized results demonstrate several important facts. First, in each domain, there are terms which are shared by speakers of all or almost all languages. These are *mountain* and *forest* in the landscape domain shared by the speakers of twelve languages, *head, hand/arm and foot/leg* in the body parts domain shared by speakers of all languages, and *cat* and *dog* in the animals domain also shared by speakers of all languages. Second, of the three domains, the body parts domain had the lowest number of unique terms (14) while the landscape domain had the highest (17). In other words, there is a higher variability in landscape domain associations than in associations in the other two domains. The variety of landscape terms reflects both the typical character of the local landscape (i.e. *sea* in the sea bordering countries) as well as the semantic association of the landscape domain with the visual appreciation and depiction of landscape in some countries (e.g. *view*, *painting*, *scenery*, *Purvītis*). Finally, when comparing the first two most cognitively salient terms in each domain, the body parts terms are 3–6 times more salient and terms for animals 2–3 times more salient than landscape terms. However, in the last two positions, landscape terms are equally or more salient than terms in the other two domains. There is therefore much disagreement on what constitutes the three domains not only between languages but also between speakers of the same language.

3.5 Semantic networks

To understand how the individual domains were structured internally, we analyzed and visualized complex semantic networks in each domain (van Putten et al. 2020; Wartmann et al. 2015). First, we created a co-occurrence matrix for each language and domain, containing counts of pairs of terms listed adjacently. This matrix was fed into Gephi (Bastian et al. 2009), a software which we used to build undirected graphs where pairs such as *hill – forest* and *forest – hill* had equal relevance for the visualized connectivity of the network. Terms were shown as nodes and co-occurrences between them as lines. Louvain community detection (Blondel et al. 2008) implemented into Gephi was used to identify related sets of terms based on their co-occurrence.

In most languages and domains, modularity values were greater than zero, meaning term clustering was not random. This was, however, not true for Croatian, Lithuanian, Polish and Ukrainian in the animal and landscape domains and for Croatian, Lithuanian and Ukrainian in the body parts domain where the modularity value was zero, i.e. there was no observable clustering. Clusters were subsequently differentiated by colors. These were

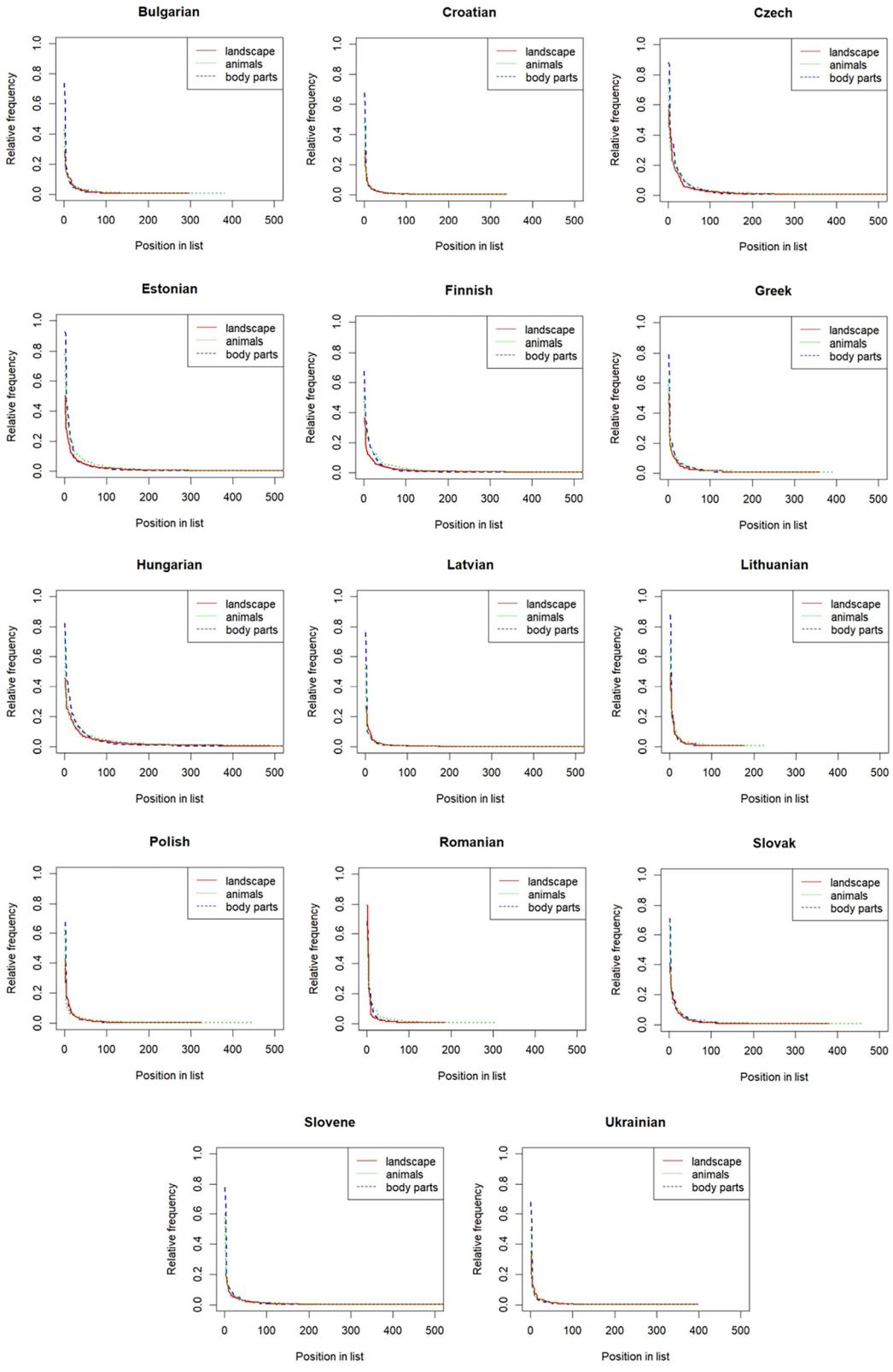

Figure 2: Frequency distribution of terms for landscape, animals and body parts. The *x*-axis corresponds to a term arranged in descending order of frequency, from most to least frequent. The *y*-axis indicates the frequency as a proportion of the total number of participants.

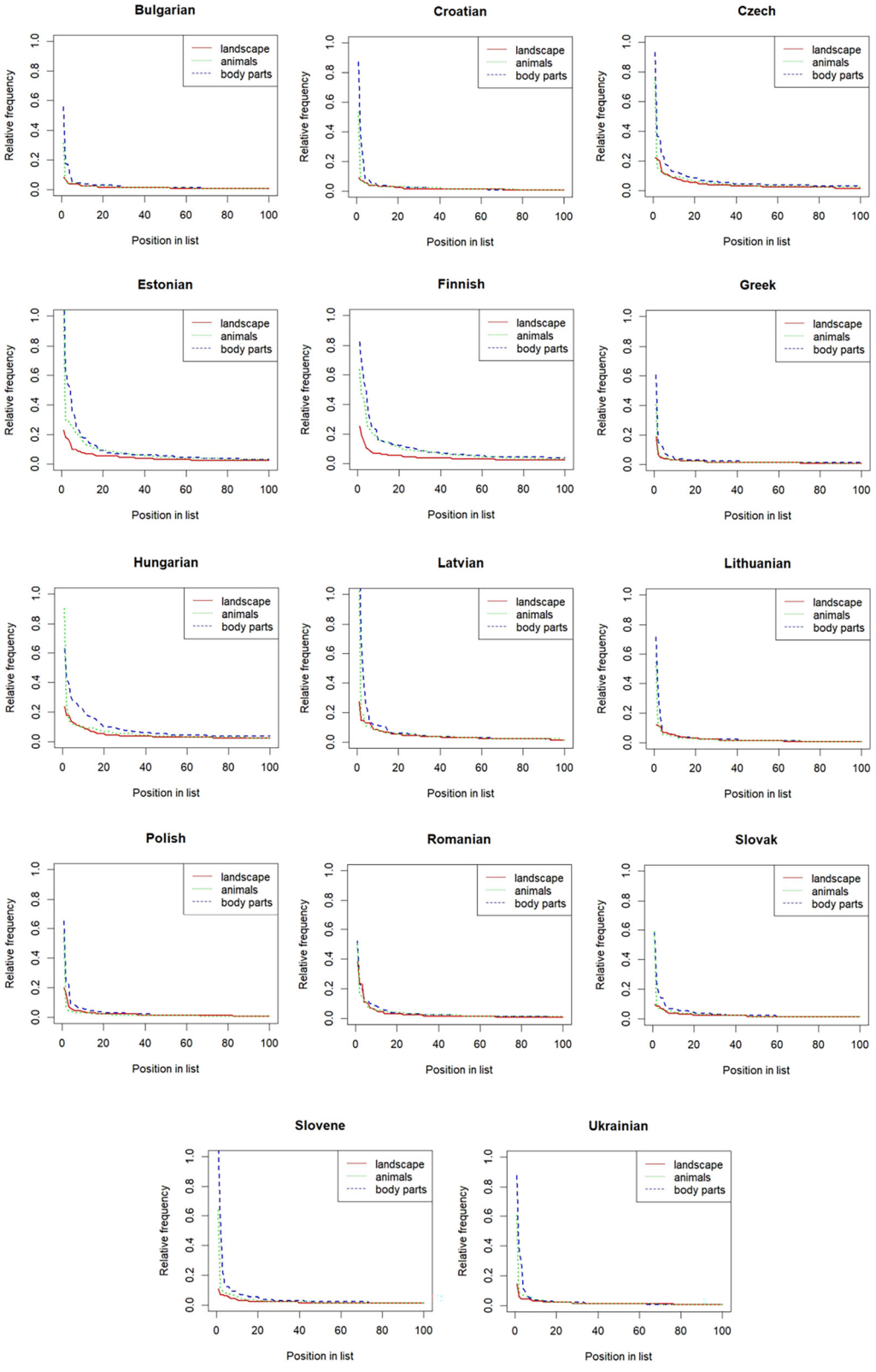

Figure 3: Frequency distribution of co-occurrence pairs for terms listed for landscape, animals and body parts domains. The *x*-axis shows the 100 most frequent co-occurrence pairs, arranged in descending order of frequency. The *y*-axis reflects their frequency as a percentage of the total participant count.

Table 3: Five most cognitively salient terms in the landscape domain by language.

Language	Term	Translation	Salience	Language	Term	Translation	Salience
Bulgarian	планина	mountain	0.08	Latvian	glezna	painting	0.10
	природа	nature	0.08		pļava	meadow	0.09
	картина	painting	0.06		daba	nature	0.09
	гора	forest	0.05		mežs	forest	0.06
	гледка	view	0.05		Purvītis	Purvītis[a]	0.05
Croatian	priroda	nature	0.13	Lithuanian	miškas	forest	0.17
	šuma	forest	0.06		kalnas	mountain	0.13
	planina	mountain	0.05		upė	river	0.10
	slika	painting	0.05		ežeras	lake	0.08
	livada	meadow	0.05		medis	tree	0.07
Czech	les	forest	0.12	Polish	góra	mountain	0.15
	louka	meadow	0.09		widok	view	0.10
	příroda	nature	0.08		las	forest	0.07
	strom	tree	0.08		morze	sea	0.06
	hora	mountain	0.08		obraz	painting	0.04
Estonian	mägi	mountain/hill	0.13	Romanian	munte	mountain	0.46
	mets	forest	0.09		deal	hill	0.23
	puu	tree	0.06		câmpie	plain	0.16
	org	valley	0.06		mare	sea	0.07
	põld	field	0.05		pădure	forest	0.06
Finnish	järvi	lake	0.10	Slovak	hora	mountain	0.11
	metsä	forest	0.09		les	forest	0.08
	meri	sea	0.07		lúka	meadow	0.07
	pelto	field	0.05		kopec	hill	0.06
	kaunis	beautiful	0.04		strom	tree	0.06
Greek	θάλασσα	sea	0.12	Slovene	narava	nature	0.07
	βουνό	mountain	0.11		gozd	forest	0.05
	φύση	nature	0.07		gora	mountain	0.05
	θέα	view	0.06		travnik	meadow	0.05
	όμορφος	beautiful	0.05		hrib	hill	0.04
Hungarian	szép	beautiful	0.08	Ukrainian	гора	mountain	0.11
	hegy	mountain	0.07		природа	nature	0.07
	fa	tree	0.06		поле	field	0.06
	erdő	forest	0.04		пейзаж	scenery	0.05
	természet	nature	0.03		море	sea	0.04

[a]Vilhelms Purvītis was a famous Latvian landscape painter and educator (1872–1945).

assigned randomly, not by semantic proximity, and their function is simply to differentiate clusters from each other within the particular language and domain, not to compare the clusters across languages.

The size of nodes reflects the frequency of occurrence and the width of edges (lines connecting nodes) reflects the frequency of co-occurrence, in both cases weighted by the calculated weighted degree (i.e. the total number of co-occurrences associated with each node). The network was then organized using the Fruchterman Reingold force-directed layout algorithm. To make the visualizations easier to interpret and to bring out the core concepts and connections in each domain, we removed all nodes with a weighted degree of less than 23.4 % of participants, following the work of van Putten et al. (2020) for easier comparison. It should be emphasized that this is a purely arbitrary threshold and alternative threshold levels may be more appropriate for different data. The visualized semantic networks are shown in Figures 4–6.

Table 4: Five most cognitively salient terms in the body parts domain by language.

Language	Term	Translation	Salience	Language	Term	Translation	Salience
Bulgarian	ръка	hand/arm	0.37	Latvian	roka	hand/arm	0.51
	крак	foot/leg	0.25		kāja	leg	0.27
	глава	head	0.23		galva	head	0.15
	крайник	limb	0.09		vēders	stomach	0.04
	торс	torso	0.05		rumpis	torso	0.02
Croatian	ruka	hand/arm	0.49	Lithuanian	ranka	hand/arm	0.42
	noga	foot/leg	0.24		koja	leg	0.33
	glava	head	0.19		galva	head	0.29
	oko	eye	0.03		pirštas	finger	0.04
	trup	torso	0.02		nosis	nose	0.03
Czech	ruka	hand/arm	0.31	Polish	ręka	hand/arm	0.33
	hlava	head	0.25		noga	foot/leg	0.26
	noha	foot/leg	0.23		głowa	head	0.20
	trup	torso	0.06		brzuch	stomach	0.04
	oko	eye	0.06		oko	eye	0.03
Estonian	käsi	hand/arm	0.38	Romanian	cap	head	0.31
	pea	head	0.37		mână	hand/arm	0.25
	jalg	foot/leg	0.31		picior	foot/leg	0.17
	kael	neck	0.07		ochi	eye	0.06
	kõht	stomach	0.06		membru	limb	0.06
Finnish	käsi	hand/arm	0.40	Slovak	ruka	hand/arm	0.28
	jalka	foot/leg	0.31		hlava	head	0.28
	pää	head	0.27		noha	foot/leg	0.19
	varvas	toe	0.04		oko	eye	0.04
	sormi	finger	0.04		telo	torso	0.04
Greek	χέρι	hand/arm	0.30	Slovene	roka	hand/arm	0.46
	πόδι	foot/leg	0.25		noga	foot/leg	0.27
	κεφάλι	head	0.21		glava	head	0.25
	μάτι	eye	0.05		telo	body	0.04
	μύτη	nose	0.04		prst	finger	0.03
Hungarian	láb	foot/leg	0.19	Ukrainian	рука	hand/arm	0.43
	fej	head	0.18		нога	foot/leg	0.25
	kéz	hand/arm	0.14		голова	head	0.20
	kar	arm	0.13		тулуб	torso	0.03
	nyak	neck	0.05		око	eye	0.02

To be able to compare networks between domains and languages, two measures were calculated: the average weighted degree and modularity. The first expresses the connectivity of a network, reflecting both the connections as well as their strength. The second measures the clustering in the network (the higher the value, the greater the clustering). Using a one-way ANOVA, we found a significant effect of the domain on average weighted degree, $F(2, 39) = 15.176$, $p < 0.001$. Pairwise comparisons with Bonferroni corrections showed significant differences between the domains of landscape and body parts ($p < 0.001$) and animals and body parts ($p = 0.005$). The difference in average weighted degree between the domains of landscape and animals was not statistically significant. Landscape showed the lowest mean value of average weighted degree ($M = 5.82$, SD = 1.1) followed by animals ($M = 6.36$, SD = 2.9) and body parts ($M = 9.88$, SD = 1.57). This measure indicates that terms in the landscape domain have fewer and weaker connections compared to the other two domains. We found no statistically significant effect of domain on modularity. The pairwise comparisons were not significant either.

Table 5: Five most cognitively salient terms in the animals domain by language.

Language	Term	Translation	Salience	Language	Term	Translation	Salience
Bulgarian	куче	dog	0.14	Latvian	suns	dog	0.27
	котка	cat	0.10		kaķis	cat	0.21
	природа	nature	0.05		lācis	bear	0.03
	див	wild	0.05		vilks	wolf	0.03
	домашен	domestic	0.04		lapsa	fox	0.03
Croatian	pas	dog	0.29	Lithuanian	šuo	dog	0.31
	mačka	cat	0.15		katė	cat	0.28
	kućni ljubimac	pet	0.06		naminiai	domestic	0.06
	divlja	wild	0.03		vilkas	wolf	0.05
	lav	lion	0.02		arklys	horse	0.04
Czech	pes	dog	0.21	Polish	pies	dog	0.28
	kočka	cat	0.16		kot	cat	0.20
	domácí	domestic	0.07		zoo	zoo	0.04
	kráva	cow	0.05		koń	horse	0.03
	divoká	wild	0.04		dzikie	wild	0.02
Estonian	koer	dog	0.24	Romanian	câine	dog	0.20
	kass	cat	0.22		pisică	cat	0.18
	lehm	cow	0.07		sălbatic	wild	0.07
	karu	bear	0.05		domestic	domestic	0.07
	hunt	wolf	0.04		cal	horse	0.06
Finnish	kissa	cat	0.26	Slovak	pes	dog	0.26
	koira	dog	0.25		mačka	cat	0.18
	hevonen	horse	0.07		kôň	horse	0.05
	lehmä	cow	0.05		domáci	pet	0.04
	lintu	bird	0.03		zoo	zoo	0.03
Greek	σκύλος	dog	0.20	Slovene	pes	dog	0.27
	γάτα	cat	0.16		maček	cat	0.21
	άγριος	wild	0.03		krava	cow	0.03
	λιοντάρι	lion	0.03		konj	horse	0.03
	άλογο	horse	0.03		divja	wild	0.02
Hungarian	kutya	dog	0.16	Ukrainian	кіт	cat	0.24
	macska	cat	0.13		собака	dog	0.21
	háziállat	domestic	0.03		корова	cow	0.03
	madár	bird	0.03		друг	friend	0.03
	ló	horse	0.03		кінь	horse	0.03

Average modularity was lowest for landscape (M = 0.13, SD = 0.11) followed by animals (M = 0.21, SD = 0.17) and body parts (M = 0.22, SD = 0.14). Compared to the other two domains, landscape terms had a lower propensity for clustering.

3.6 Individual differences in landscape experience

We followed van Putten et al. also to see if the length of participants' outdoor activities and the frequency of personal experience with various landscape objects correlated with the number of terms free-listed in participants' responses. Calculating Spearman's rho correlation coefficient, we found a weak positive correlation between the amount of time spent in outdoor activities and the number of listed terms, $r_S(2{,}970) = 0.11$, $p < 0.001$. There was no statistically significant correlation between the number of listed objects and landscape objects visited at least once a month, $r_S(2{,}970) = 0.01$, $p = 0.550$, or landscape objects ever visited, $r_S(2{,}970) = -0.01$, $p = 0.795$.

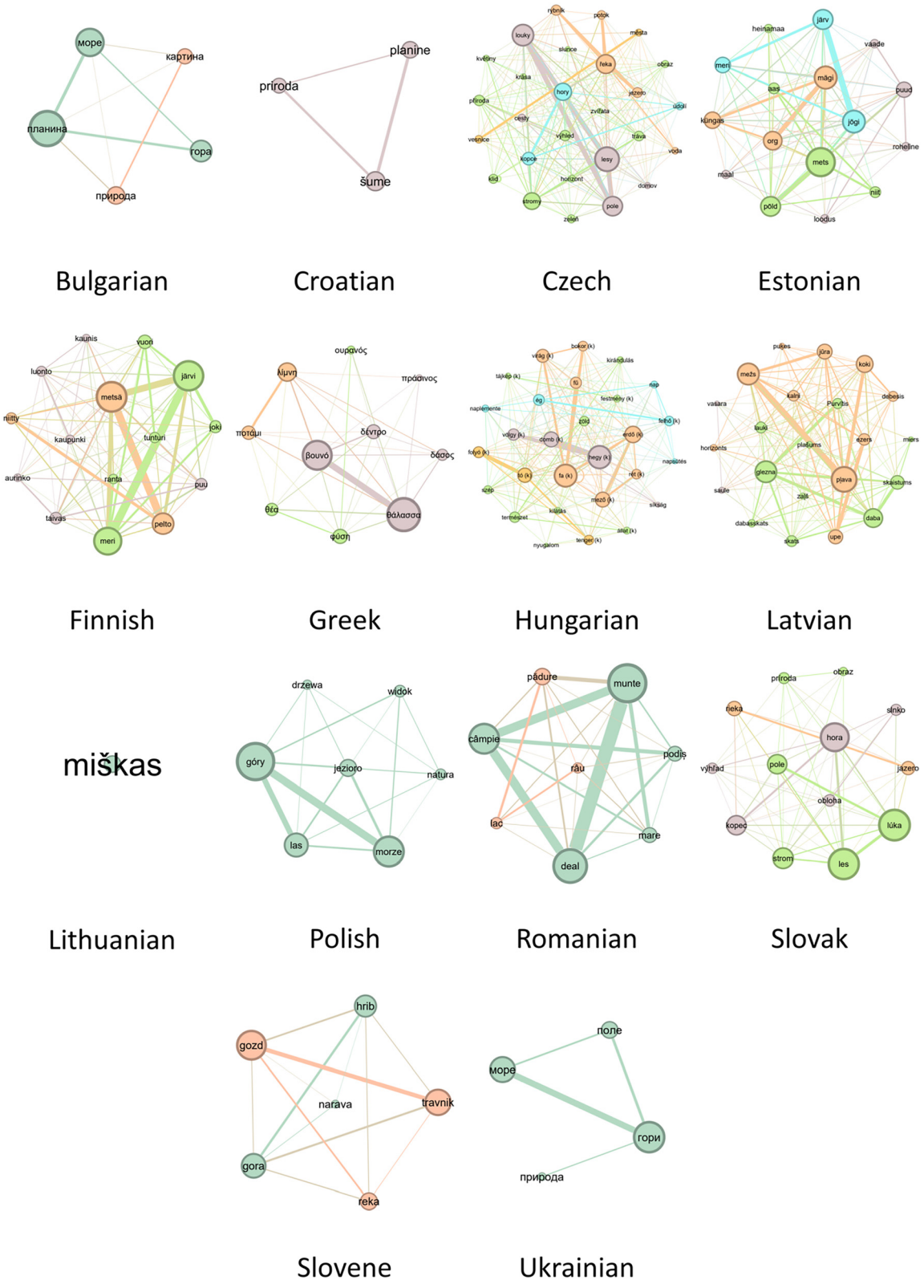

Figure 4: Semantic networks in the landscape domain by language. Nodes represent terms, links signify co-occurrence, and colors differentiate clusters within each language.

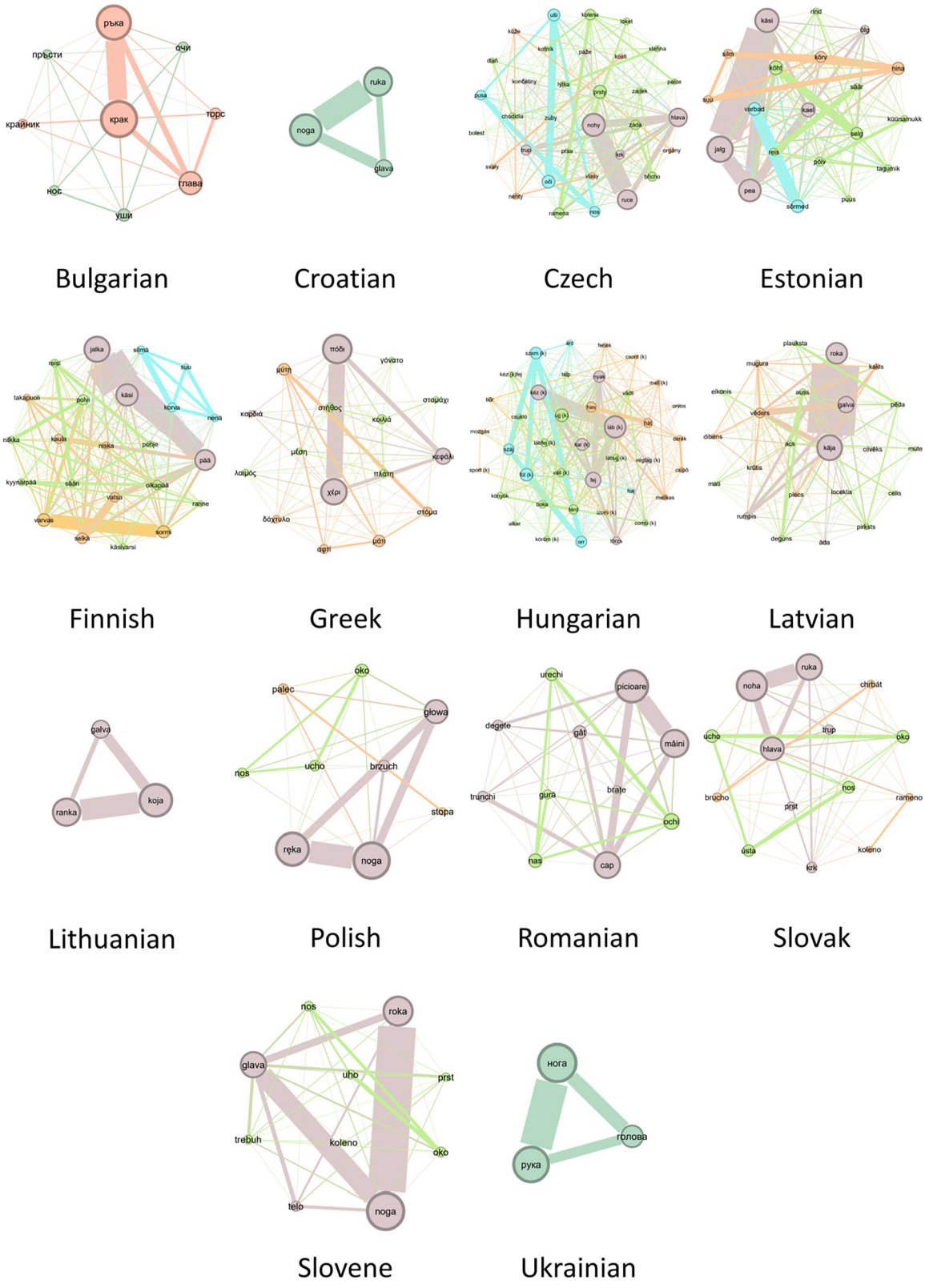

Figure 5: Semantic networks in the body parts domain by language. Nodes represent terms, links signify co-occurrence, and colors differentiate clusters within each language.

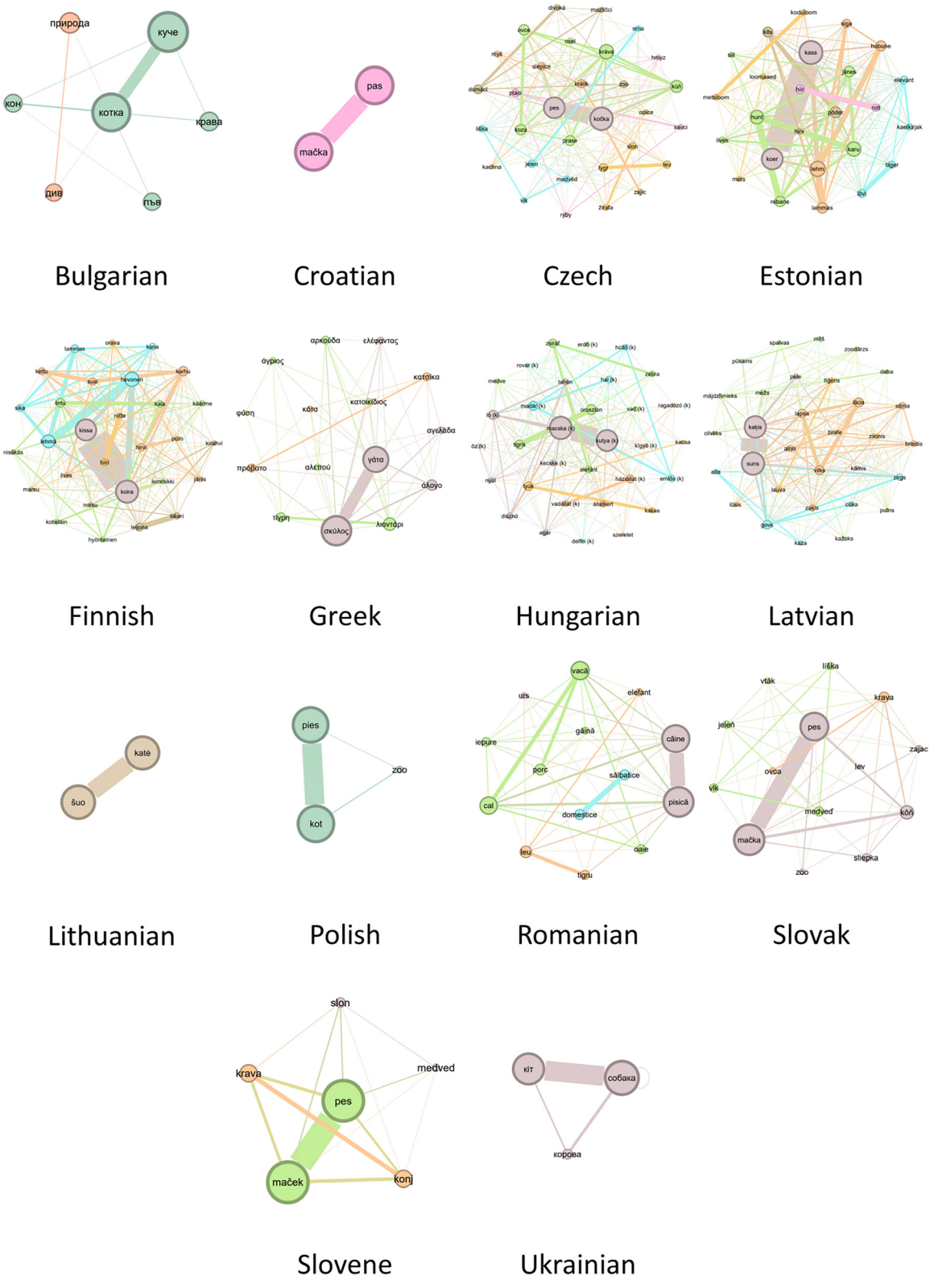

Figure 6: Semantic networks in the animals domain by language. Nodes represent terms, links signify co-occurrence, and colors differentiate clusters within each language.

4 Discussion

In general, the participants of our research showed similar linguistic characteristics as were reported in van Putten et al.'s study (2020). The landscape domain seems to be less well-defined than the animals and body parts domains. Subjectively, the participants considered listing terms in the landscape domain as more difficult than listing terms in the other two domains. This was reflected in the fact that they listed fewer landscape terms, these terms were listed with the lowest frequency and they had the smallest proportion of shared co-occurrence pairs. Furthermore, there was a broader variety of the most cognitively salient terms in the landscape domain than in the other two domains. All of this suggests that the landscape domain may be more weakly structured than the other two domains at both the *inter*-language as well as the *intra*-language level.

We also have to concur that the internal structuring of the landscape domain varied across languages as demonstrated by the cognitive salience and semantic networks analyses. For example, *mountain* appears to represent landscape across European languages, just as van Putten et al. found in their study. However, we were not able to identify any patterns characteristic of language families similar to those suggested by van Putten et al. *Sea*, for example, played a prominent role in (some) sea-bordering countries, but it was ostentatiously absent in Croatian, Slovene and Estonian. In the first two cases, this is probably the result of the residential composition of the research sample (predominance of Zagreb and Ljubljana residents), reflecting the local experience of landscape. We would expect to find more references to the sea if this research were done on the coast. This would indicate the importance of environmental factors in landscape conceptualization as suggested by Hansen and O'Meara (2020). Furthermore, in the Slovene national imagination, mountains, rather than the sea, occupy a more prominent role, so cultural factors may be at play as well. On the other hand, in Estonia, the low frequency of references to the sea (and sea-related landscape features) in free-listed responses may be the direct result of the question formulation. The word *maastik* was used as a translation for landscape. Estonian, however, has a separate word for seascape (*marevaade*), associated with a partially different cognitive and linguistic domain. As a consequence, the question formulation may have limited the range of responses we were able to collect. This would, by contrast, indicate the importance of language in landscape conceptualization.

More apparent was the clustering by semantic proximity. For example, water-related terms often formed distinct clusters (Czech, Estonian, Finnish, Hungarian, Slovak) as did characteristic, rhythmically changing features of the local cultural landscape such as *field*, *meadow* and *forest* (Czech, Estonian, Finnish, Slovak) or geomorphologically related terms such as *mountain*, *hill* and *valley* (Czech, Estonian, Hungarian, Slovak, Slovene). Another very interesting result is the aesthetic and visual understanding of landscape as evidenced by terms such as *beautiful*, *view*, *scenery*, *painting* and *Purvītis*. One cannot but recall the seminal work of Denis Cosgrove in which he defined landscape as a historically conditioned way of seeing (re-)produced by hegemonic power relations that has to be learned (i.e. landscape as a visual ideology) rather than as an external physical environment naturally existing independent of perception (Cosgrove 1998). This would also explain why terms such as *mountain* and *forest* seem to symbolize landscape in nearly all languages and even in countries which are very flat such as Lithuania and Estonia (or the Netherlands in van Putten et al.'s data) or countries with proportionately little forest cover such as Hungary. To a significant extent, Europeans have learned to perceive the landscape in similar ways despite their linguistic differences. Rather than as a piece of strong evidence for linguistic relativity, we, therefore, see these results as a reflection of the diversity of local landscapes, the ideological dimension of landscape portrayal and aesthetics, and the internal semantic organization of the landscape domain. Landscape perception seems to be the result of complex factors of which language is but one. A holistic approach is therefore needed (Antrop and van Eetvelde 2017).

Furthermore, a detailed analysis of our results shows a somewhat more complicated picture than was reported by van Putten et al. Differences in difficulty, number of listed terms and co-occurrence pairs, degree of cognitive salience and other trends were not identified in all languages equally. Even where they were observed, in many cases they were not statistically significant. This may be due to methodological limits (see below) but it may also be an indication of a lack of fundamental differences between languages or the influence of extra-linguistic factors for landscape conceptualization. For example, despite identical methodology, we did not

observe any preference for celestial or water-related phenomena in Romanian which van Putten et al. identified in their sample of Romance languages and which would set Romanian apart from the other languages and language families in our sample. This may be the result of the linguistic and/or cultural influence of surrounding Slavic and Hungarian nations but it may also be a potential test for the validity of van Putten et al.'s claim for language family-level differences in landscape conceptualization. By contrast, *mountain* and *sea* co-occurred in the same association frequently (Bulgarian, Greek, Polish, Romanian, Ukrainian) across language families and the only language with a clear association of *mountain* with celestial phenomena was Slovak. Further research is, therefore, clearly needed.

To do this, we need to engage in a critical discussion about appropriate methodology, identify the limits of applied methods, and suggest alternative courses of research. Comparing our results with van Putten et al.'s, it is apparent that although we had a larger sample overall and for each language, our participants listed, on average, three times fewer responses. The total number of provided terms differed between languages with speakers of Hungarian and Czech giving the highest average number of landscape terms (18 and 16) and speakers of Croatian and Latvian offering the least (4 and 5). In addition, in some languages (namely Croatian, Latvian, Slovene and Ukrainian), the proportion of one-word answers was relatively high (18–35 %) but even discounting one-word answers, the average number of listed terms in most languages was still low and only four languages exceeded 10 (Czech, Estonian, Finnish and Hungarian). This may have had a negative effect on the statistical significance of calculations of frequency of occurrence and co-occurrence pairs, the estimation of cognitive salience and the construction of semantic networks even though our results were generally aligned with van Putten et al.

We do not have a clear explanation for this problem because instructions for respondents in all languages were identical. The form was distributed on-line, the respondents filled it out independently, and the researchers could not influence respondents by any additional instructions. We speculate that the low number of responses in some languages may have been due to the more frequent use of mobile phones rather than computers for answering the form. Small screens, small keyboards, and the "scrolling mentality" of phone use may have led to less patience with the form and a lower willingness to use the three-minute limit in its entirety. We recommend that future researchers pay extra attention to encouraging participants to list as many responses as possible within the given time. Taking growing mobile phone use for on-line surveys as a given, one possible solution to the problem of low response rate could be to implement audio input into the form, so respondents can simply dictate free lists instead of typing them.

While this may have been a problem that we introduced into the research, we believe other aspects of the research should be re-evaluated for future uses as well. First, as noted above and as commented on by van Putten et al. for their study, the creation of forms in different languages presents a double challenge. All participants across languages must receive identical prompts which is a common problem in international social science research. This research, however, brought an additional challenge arising from its research object: language. There is a high risk of tautology in this process. The choice of terms is hence a highly delicate matter significantly influencing participants' responses. By constructing a form about landscape, we run the risk of constructing the landscape we want to learn about.

As we pointed out above, it was not only difficult to translate the landscape domain itself but also the individual landscape terms appearing in the form. Greek, for example, does not have a single word for *coast*. The differentiation of *mountains (higher than 1,000 m)* and *hills (lower than 1,000 m)* by van Putten et al. was a self-fulfilling prophecy not corresponding to the differentiation of mountains, hills and other convex landscape features in individual languages or the mountain-hiking experience of their speakers. By contrast, Estonian only has one word (*mägi*) for both mountains and hills. Czech has a word for *lake* but there are almost no lakes in Czechia because the term is used only for naturally formed water bodies, unlike ponds and dams which are plentiful, intimately known and forming entire emblematic cultural landscapes (but which were not offered in the form although participants listed them twice as frequently as lakes). In other words, the form contained terms often irrelevant for many participants because of their lack in the local landscape (such as volcano and desert) or their lack in the local language (e.g. the case of mountains and hills in Estonian). On the other hand, potentially more salient terms which could be more productively correlated with the free-listed responses to see if physical experience with landscape objects correlated with landscape conceptualization were missing. This could explain

the absence of a relevant correlation between landscape experience and free-listed responses in both our and van Putten et al.'s data.

Analogically, it was very complicated to translate participants' responses back into English without losing the subtle nuances of meaning (or mixing up meanings in the case of polysemic terms). Words would rarely have one-to-one translations. The nominal equivalence of terms may not correspond with the rich meaning(s) they have in everyday use. Standardization which is necessary for statistical analyses may erase such differences and create a false picture of the semantic organization of the linguistic domain. For example, the unification of singular and plural forms may have obfuscated differences in landscape perception in terms of scale and understanding. In each language, some words appeared predominantly in the plural while others appeared in the singular. Is it relevant? Finally, spelling differences may betray differences in landscape understanding in languages with a high degree of dialectal variations.

One potential solution to the risk of tautology is to use a two-step process. First, ask one sample of participants to free-list terms associated with a certain domain, create a list of the most salient landscape objects for each language, and then ask a second group of participants to free-list responses and choose objects from the list created with the first group to see if experience with specific landscape features corresponds to listed terms. Of course, this is only possible for those languages that have a term for landscape and differentiate the landscape domain from other domains. If this is not the case, such as in Tzeltal (Brown 2008), the process has to begin with the participants' delimitation of the domain for which we want to generate lists of terms.

Still, this may not be sufficient to avoid ambiguity and polysemy. From lists of terms alone we have no way of knowing if participants understand the terms identically even within a single language, let alone across many languages. It may be that their images of landscape may be similar even though they use different terms to describe them (and vice versa). The correct procedure would then be to incorporate yet another step into the research process in which participants would be asked to describe, draw or photograph instances of listed terms. This would add quality to the data and with sufficient numbers a statistical analysis of such descriptions, drawings and photographs could also be carried out. We therefore support the call for the use of micro-ethnographies to understand better the localized conceptualizations of landscape (Downey and Gillett 2023).

5 Conclusions

The landscape domain seems to be more loosely structured than the animals and body parts domains. Language speakers find the landscape domain more difficult to define and there is less consensus among them on what constitutes it. We also found some variability in landscape conceptualizations across languages but this variability does not seem to be systematic. At this moment, we are unable to conclude that speakers of different languages conceptualize landscape differently on a fundamental level. Multiple factors influencing landscape perception appear to be at play and landscape conceptualizations seem to differ individually, geographically, linguistically, as well as culturally. We must therefore strongly support van Putten et al.'s call to incorporate local understandings of the landscape into regional, national, and European landscape policies. Research such as van Putten et al.'s and ours helps to break the path for such an understanding.

In relation to the linguistic relativity debate, a more complex methodology has to be designed to test for the effect of language on landscape perception. Recent research suggests that speakers of different languages may differ in landscape perception in their sensory, motor, and emotional associations of individual landscape categories (Striedl et al. 2024). However, isolating the precise influence of linguistic and extra-linguistic factors and incorporating local qualitative understandings of landscape terms remains a challenge.

While a unified methodology provides valuable insights by highlighting both similarities and differences across languages and language families, our overall contribution to the linguistic relativity hypothesis may not be as substantial as initially anticipated due to apparent cultural commonalities. The languages we examined all originate from cultures that Majid and Levinson (2010) called WEIRD (Western, Educated, Industrialized, Rich, and Democratic). This shared cultural and socioeconomic background may increase the likelihood of similarities in conceptualizations, outweighing potential linguistic differences. Expanding the comparative analysis to

include non-European languages, cultures and geographies such as Native North America (e.g. Basso 1996; O'Neill 2008), Aboriginal Australia (e.g. Koch and Hercus 2009) or Africa (e.g. Krohmer 2010) is therefore a must for future research.

Funding information: The work was financed by 68378076 (RVO), Czech Academy of Sciences (Přemysl Mácha), P6-0101, Slovenian Research Agency (Matjaž Geršič), MK, 2/0120/24, VEGA, Slovakia (Ján Mačutek), APVV-21-0216, Slovak Research and Development Agency (Michaela Koščová).
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission. All authors participated in research design, all authors with the exception of JM and MK participated in data collection. All authors participated in data standardization. JM and MK were responsible for the statistical analysis. PM prepared the manuscript with contributions from all co-authors.
Conflict of interest: Authors state no conflict of interest.
Data availability statement: The datasets generated and analyzed during the study, including the PDF versions of questionnaires in English and all 14 research languages, are available for full download at the Data Repository of the Czech Academy of Science. DOI: https://doi.org/10.57680/asep.0618765.

References

Antrop, Marc & Veerle van Eetvelde. 2017. *Landscape perspectives: The holistic nature of landscape*. Dordrecht, Netherlands: Springer.
Baier, Diane, Soonja Choi, Florian Goller, Yunju Nam & Ansorge Ulrich. 2023. Does language rule perception? Testing a radical view of linguistic relativity. *Journal of Experimental Psychology: General* 152(3). 794–824.
Basso, Keith H. 1996. *Wisdom sits in places: Landscape and language among the Western Apache*. Albuquerque: UNM Press.
Bastian, Mathieu, Sebastien Heymann & Mathieu Jacomy. 2009. Gephi: An open source software for exploring and manipulating networks. In *Proceedings of the International AAAI Conference on Web and Social Media* vol. 3, 361–362. Available at: http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154.
Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte & Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 10. P10008.
Brown, Penelope. 2008. Up, down, and across the land: Landscape terms, place names, and spatial language in Tzeltal. *Language Sciences* 30(2–3). 151–181.
Cosgrove, Denis E. 1998. *Social formation and symbolic landscape*. Madison: University of Wisconsin Press.
Danesi, Marcel. 2021. *Linguistic relativity today: Language, mind, society, and the foundations of linguistic anthropology*. New York: Routledge.
Downey, Greg & Alexander J. Gillett. 2023. Linguistic relativity in cross-cultural context: Converging evidence from neuroanthropology. *Topics in Cognitive Science* 15(4). 693–697.
Hansen, Magnus Pharao & Carolyn O'Meara. 2020. Environmental adaptation in language: Spatial grammar, landscape knowledge and human survival. *Language Dynamics and Change* 10(2). 230–258.
Haun, Daniel B. M., Christian J. Rapold, Gabriele Janzen & Stephen C. Levinson. 2011. Plasticity of human spatial cognition: Spatial language and cognition covary across cultures. *Cognition* 119(1). 70–80.
Heegård, Jan & Henrik Liljegren. 2018. Geomorphic coding in Palula and Kalasha. *Acta Linguistica Hafniensia* 50(2). 129–160.
Johnson, Leslie M. & Eugene S. Hunn (eds.). 2010. *Landscape ethnoecology: Concepts of biotic and physical space*. New York – Oxford: Berghahn Books.
Koch, Harold & Luise Hercus. 2009. *Aboriginal placenames: Naming and re-naming the Australian landscape*. Canberra: ANU Press.
Krohmer, Julia. 2010. Landscape perception, classification, and use among Sahelian Fulani in Burkina Faso. In Leslie M. Johnson & Eugene S. Hunn (eds.), *Landscape ethnoecology: Concepts of biotic and physical space*, 49–82. New York – Oxford: Berghahn Books.
Levinson, Stephen C. 1996. Language and space. *Annual Review of Anthropology* 25(1). 353–382.
Levinson, Stephen C. 2003. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.
Levinson, Stephen C. & David P. Wilkins (eds.). 2006. *Grammars of space: Explorations in cognitive diversity*, vol. 6. Cambridge: Cambridge University Press.
Lucy, John A. 2016. Recent advances in the study of linguistic relativity in historical context: A critical assessment. *Language Learning* 66(3). 487–515.
Majid, Asifa & Stephen C. Levinson. 2010. WEIRD languages have misled us, too. *Behavioral and Brain Sciences* 33(2–3). 103.
Malt, Barbara C. & Asifa Majid. 2013. How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science* 4(6). 583–597.
Mark, David M., Andrew G. Turk, Niclas Burenhult & David Stea (eds.). 2011. *Landscape in language: Transdisciplinary perspectives*, 4. Amsterdam/Philadelphia: John Benjamins Publishing.
O'Neill, Sean. 2008. *Cultural contact and linguistic relativity among the Indians of northwestern California*. Norman: University of Oklahoma Press.

Striedl, Philipp, Asifa Majid & Ross S. Purves. 2024. Sensory, motor, and emotion associations for landscape concepts differ across neighbouring speech communities. *Landscape Research* 49(5). 704–721.

Sutrop, Urmas. 2001. List task and a cognitive salience index. *Field Methods* 13. 263–276.

Urban, Matthias. 2020. Mountain linguistics. *Language and Linguistics Compass* 14(9). e12393.

van Putten, Saskia, Carolyn O'Meara, Flurina Wartmann, Joanne Yager, Julia Villette, Claudia Mazzuca, Claudia Bieling, Niclas Burenhult, Ross Purves & Asifa Majid. 2020. Conceptualisations of landscape differ across European languages. *PLoS One* 15(10). e0239858.

Wartmann, Flurina M., Ekaterina Egorova, Curdin Derungs, David M. Mark & Ross S. Purves. 2015. More than a list: What outdoor free listings of landscape categories reveal about commonsense geographic concepts and memory search strategies. In Sara Irina Fabrikant, Martin Raubal, Michela Bertolotto, Clare Davies, Scott Freundschuh & Scott Bell (eds.), *Spatial information theory*, 224–243. Cham: Springer International Publishing.

Wolff, Phillip & Kevin J. Holmes. 2011. Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science* 2(3). 253–265.