The differential coding of perception in the world's languages

Asifa Majid[1,2*], Seán G. Roberts[4], Ludy Cilissen[1],

Karen Emmorey[5], Brenda Nicodemus[6], Lucinda O'Grady[5], Bencie Woll[7], Barbara LeLan[8],

Hilário de Sousa[1], Brian L. Cansler[9], Shakila Shayan[10], Connie de Vos[1,2], Gunter Senft[1],

N. J. Enfield[11], Rogayah A. Razak[12], Sebastian Fedden[13], Sylvia Tufvesson[1],

Mark Dingemanse[1], Ozge Ozturk[14], Penelope Brown[1], Clair Hill[1,11,15], Olivier Le Guen[16],

Vincent Hirtzel[17], Rik van Gijn[18], Mark A. Sicoli[19], & Stephen C. Levinson[1,2*]

[1] Max Planck Institute for Psycholinguistics, The Netherlands
[2] Radboud University, The Netherlands
[4] University of Bristol, United Kingdom
[5] San Diego State University, USA
[6] Gallaudet University, USA
[7] University College London, United Kingdom
[8] Université Paris IV – Sorbonne, France
[9] University of North Carolina at Chapel Hill, USA
[10] Utrecht University, The Netherlands
[11] The University of Sydney, Australia
[12] Universiti Kebangsaan Malaysia, Malaysia
[13] Université Paris 3 (Sorbonne-Nouvelle), France
[14] University of Cambridge, United Kingdom
[15] Lund University, Sweden
[16] CIESAS, Mexico
[17] CNRS/Paris Nanterre University, France
[18] University of Zurich, Switzerland
[19] University of Virginia, USA

* Corresponding authors: asifa.majid@let.ru.nl stephen.levinson@mpi.nl

SOCIAL SCIENCES: Psychological and Cognitive Sciences; Anthropology

Abstract

Is there a universal hierarchy of the senses, such that some senses (e.g., vision) are more accessible to consciousness and linguistic description than others (e.g., smell)? The long-standing presumption in Western thought has been that vision and audition are more objective than the other senses, serving as the basis of knowledge and understanding; whereas touch, taste, and smell are crude and of little value. This predicts that humans ought to be better at communicating about sight and hearing than the other senses, and decades of work based on English and related languages certainly suggests this is true. But how well does this reflect the diversity of languages and communities worldwide? In order to test whether there is a universal hierarchy of the senses, stimuli from the five basic senses were used to elicit descriptions in twenty diverse languages, including three unrelated sign languages. We found that languages differ fundamentally in which sensory domains they linguistically code systematically, and how they do so. The tendency for better coding in some domains can be explained in part by cultural pre-occupations. Although languages seem free to elaborate specific sensory domains, some general tendencies emerge: e.g., with some exceptions, smell is poorly coded. The surprise is that despite the gradual phylogenetic accumulation of the senses, and the imbalances in the neural tissue dedicated to them, no single hierarchy of the senses imposes itself upon language.

Significance Statement

It has long been thought, following Aristotle, that the distal senses (vision, audition) have primacy over the proximal senses (touch, taste, smell): the distal ones are after all served by more evolutionarily advanced organs and more brain tissue, and appear more accessible to consciousness and precise linguistic description. But in this study of 20 diverse languages we show this apparent primacy is by no means universal: In many languages taste outranks vision for linguistic codability, auditory experience is surprisingly ineffable in many languages, and smell although poorly coded cross-culturally nevertheless can be linguistically coded. These surprising results suggest that cultures may differ fundamentally in the accent they put on different senses as reflected in our abilities to talk about them.

There are few more compelling questions in cross-cultural research than whether other people perceive the world the way we do. Language has the potential to offer insight on the matter, and a great deal of cross-linguistic research has taken as its litmus test the nature of color terms, which appear to vary across cultures, but within constraints (1–6). Some languages, for example, make a distinction between “light blue” and “dark blue”, and there are correlated perceptual consequences of distinct linguistic categories (7–12). Despite these effects of language on perception, it is clear that perception is partly independent of language: Indeed, language seems to have distinct limitations on coding in certain domains. For example, English provides terms for simple geometric shapes (circles, squares, triangles, etc.), but describing a face so that it can be recognized is extremely challenging; similarly colors can be named with relative ease, but smells seem to resist precise description.

Since Aristotle, it has been supposed that there is a hierarchy of the senses with sight the dominant sense, followed by hearing, smell, touch and taste (2), opening the possibility that some aspects of perception are intrinsically more accessible to consciousness and thus to language. The position of smell has since been further demoted, based in part on insights regarding the evolutionary development of our senses, in which stereoscopic vision, wider eye orbits and increasing visual cortex have evolved at the expense of the olfactory bulb and olfactory epithelium (13–19). Modern re-workings of the Aristotelian hierarchy give primacy to sight followed by hearing, touch, and then taste and smell (20, 21). Regardless of the precise characterization, the distal senses of vision and audition are privileged at the expense of the lowly proximal senses of touch, taste, and smell (22).

The idea that differential expressability (or conversely, ineffability) might tell us something specific about the innate architecture of cognition and how the different faculties can “talk” to language is extremely attractive. But it rests on the presumption that these patterns are universal, and invariant across languages and cultures. Many scholars, for example, have opined that no language will have a developed lexicon for smell (23–28). This presumption has been challenged by cross-cultural investigations showing that while English-speakers may indeed display the oft-touted visual dominance, other cultures show a different picture altogether (29). For example, the Jahai—a hunter-gatherer community residing in the rainforests of the Malay Peninsula—find odors just as easy to express as visual entities (30). This raises the question of whether the hierarchy of the senses is universally expressed across languages.

To test this, we explored the coding of percepts involving different senses (vision, hearing, touch, taste, smell) in 20 largely unrelated languages around the world. The languages were sampled to reflect linguistic diversity from each of the major landmasses, drawn from 15 distinct language families (Figure 1, Table 1). Many of these languages are spoken by small ethnic groups with distinctive cultures of their own. We included three unrelated sign languages, allowing us to explore the influence of the modality in which language is expressed (vocal or gestural). Obviously for the deaf native signers the coding of auditory stimuli had to be omitted in this comparative study.

We explored the 5 perceptual domains, and investigated both color and shape for vision. Color, of course, has a long history of cross-cultural exploration (1, 31); we added shape in our battery of tests in order to examine whether the predicted primacy of vision generalized across sub-domains. Color stimuli were sampled from Munsell color space, basic geometric figures explored shape distinctions, auditory stimuli varied in pitch, loudness and tempo, tactile stimuli focused on tactile texture distinctions (e.g., rough, smooth), the basic taste distinctions of sweet, salty, sour, bitter and umami were sampled, and smell was tested using a standard olfactory stimulus of micro-encapsulated odors depicting common scents (e.g., onion, smoke). It should be borne in mind that these domains have some intrinsic differences: for example, the psychophysical space for color is three-dimensional, but odor has an unknown multitude of dimensions. Consequently, any conclusions regarding the absolute codability of domains have to take this fact into consideration. However, our main question concerns whether languages differ in their coding of the senses, specifically in the relative expressibility of perceptual domains. If there is variation across languages in how domains encode the senses, this would not be easily explained by intrinsic domain constraints. We return to this issue in the discussion.

All of the language data were collected by fieldworkers committed to long-term documentation of these languages, or by researchers otherwise experts in the languages (Figure 1, Table 1) using a standardized "field manual" (8; see Methods). Collaborative field work meant that skilled researchers were fluent in the local languages, and because of their long-term engagement with communities were able to recruit participants in their local habitat. Testing was conducted in local languages, with fieldworkers asking participants to name stimuli in each of the sensory domains with the same protocol. Researchers then transcribed responses according to a standardized coding scheme, and identified the main semantic elements which expressed the perceptual domain at hand. For example, for the full response *light moss green* given to a color chip, the main contentful response was coded as *green*. The same procedure was used across modalities. The coding process was iterated across languages until we had standardized the procedure. This was necessary because what is coded as distinct words in English, for example, can be coded in a single word in another language with morphology. Take Tzeltal, *tzajtzajtik*—a single response to a red stimulus—it is actually morphologically complex, being made up of the root morpheme for red *tzaj*, reduplicated with the suffix *tik* a morpheme which maintains the adjectival status of the root morpheme, 'red-red-AJ', i.e., sort of red (33). This is another reason why a collaborative endeavor with expert field linguists is necessary for such a project: it optimizes both standardization for language comparison (i.e., making sure languages are coded in equivalent ways), while doing justice to each language's particulars without carelessly glossing over critical differences. Once all tricky cases had been identified and appropriately treated, a uniform protocol was applied and checked across all languages (SI Appendix, S1).

Results

Relative codability of the senses

We asked whether the senses are equally expressible (or, alternatively, ineffable) in all languages. Earlier work in the color domain has operationalized this notion by referring to "codability": Brown and Lenneberg (34) showed for example that length of response (number of syllables or words), reaction times, agreement across speakers and within speakers over time, all correlated highly, but that agreement across speakers had by far the highest factor loading. They also showed that codability correlated with correct recognition of colors (see also 13), concluding that "more nameable categories are nearer the top of the cognitive 'deck'" (34). Codability is thus an important measure, rolling in Zipf's law (frequency of names correlates with brevity) with perceptual accessibility (36).

Following this tradition, we took as our operational definition of codability the degree to which a stimulus was consistently named within a language community. A measure that reflects this is Simpson's Diversity Index D (37), borrowed from ecology (where it is used to measure species diversity taking into account both the types and abundance of species). This measure has been used previously in language comparison, where naming diversity is calculated taking into account both the type and frequency of labels per stimulus (30, 38). For a given stimulus within a language, if speakers produce N description tokens, including R unique description types from 1 to R, each with frequencies of n_1 to n_R, then Simpson's Diversity Index is:

$$D = \frac{\sum_{i=1}^{R} n_i(n_i - 1)}{N(N-1)}$$

An index of 1 indicates high codability (and low naming diversity): all respondents produced the same description; whereas 0 indicates low codability (conversely high naming diversity) since all respondents produced different unique descriptions (note: absence of an overt description was treated as a unique type in the following analyses). For each stimulus in each domain, we calculated codability per stimulus per language community, and then compared these values.

Other measures of agreement are also possible, of course; such as the Shannon information index (39) and the "interpersonal agreement" measure used in the classic work of Brown and Lenneberg (34), but these do not adjust for the number of responses from a particular community. In any case, in our sample the different measures were highly correlated (Shannon, $r = -0.97$; Brown~Lenneberg interpersonal agreement, $r = 0.95$, see SI Appendix, S3).

We used mixed effects modeling in R (40, 41) to test whether there is a universal hierarchy of the senses or, alternatively, whether languages differed. The full model included random intercepts for stimulus, domain, language, and the interaction between language and domain. Log-likelihood comparison was used to compare the full model to a model without one of those intercepts (see SI Appendix, S4 for further details). The full dataset contained 44,091 descriptions from 313 respondents in 20 languages. We found languages differed in codability ($\chi^2 = 4.6$, $p = .03$), as did perceptual domains ($\chi^2 = 27.8$, $p < .001$). Crucially, however, codability for domains differed across languages ($\chi^2 = 700.0$, $p < .001$), meaning there is no universal hierarchy of the senses (Figure 2).

A skeptic might argue that the variation we see across languages is mere noise around a universal pattern. If so, a closer examination of the main effect of domain might reveal the pan-human hierarchy of the senses. If you seek a single hierarchy of the senses that generalizes over the whole language sample using a decision tree (clustering) with random effects for language and stimulus type, the following order emerges (from most to least codable): [color, taste] > [shape, sound, touch] > smell, and a permutation test confirmed the same broad pattern: [color, taste] > shape > [sound, touch] > smell (SI Appendix, S4. This was also supported by a Skillings-Mack test on rankings; SI Appendix, S5). Whatever the precise position of shape in this ordering, it is nevertheless clear that this overall cross-linguistic hierarchy is not the widely presumed Aristotelian one.

This generalized ranking of the senses across languages does not do justice to the attested cross-cultural variation in the hierarchy of the senses. In fact, out of 20 languages there are 13 unique rankings of perceptual domain by mean codability (see Figure 3). In Malay, for example, shape is the most codable of the senses on average and smell is the least codable, but in Umpila the exact opposite pattern holds—smell is the most codable and shape the least. The attested unique rankings are fewer than would be expected by totally random sampling (permutation $z = -6.5$, $p < 0.001$), suggesting that the ranking of the senses is not entirely arbitrary, as also suggested by the decision tree analysis above. Across this diverse sample no language had a domain ranking compatible with the predicted Aristotelian order (sight > sound > touch > taste > smell; see Figure 3). The closest to this hierarchy was English, which application of Spearman's footrule showed was closer to the predicted order than would be expected by chance ($p = 0.01$; see SI Appendix, S5). All other languages showed no greater fit to the Aristotelian order than would be expected by chance (SI Appendix, S5).

As discussed in the introduction, the senses differ in their inherent psychophysical dimensionality. Color is three-dimensional, taste arguably four-dimensional (42, 43), sound (leaving aside timing and timbre for a moment) is two-dimensional (44, 45); whereas shape, touch and smell likely vary on many more dimensions. The broad regularities noted above cannot therefore follow directly from the dimensionality of psychophysical spaces. A corollary could be that the hierarchy stems from differences in stimulus sampling (see Methods): perhaps presenting fewer stimuli in a domain leads to higher estimated codability because people can focus better on the task; or conversely presenting more stimuli leads to higher estimated codability because it is less likely to be skewed by an aberrant datapoint. But we found no relation between number of stimuli used in the experiment and the resulting codability attested. We sampled 80 distinct colors, but only 5 tastants; and yet exactly these two domains (color, taste) showed equal mean codability (collapsing across languages). An explicit test of whether the number of stimuli in the experiment predicted codability finds no support ($\chi^2 = 1.1$, $p = 0.28$; SI Appendix, S4).

Our sample included both signed and spoken languages, so we can specifically examine whether modality of language affected codability of perceptual domains. For example, it is widely held that there is a trade-off between the senses, such that loss of one perceptual sense leads to heightened abilities in the other senses (46). This might predict that sign languages would show higher codability for non-auditory modalities. To assess whether signed languages clustered together, we used regression trees with random effects for each stimulus (47) to predict codability by language, domain, and modality. There were no simple generalizations that classified signed languages as distinct from spoken languages; i.e., modality of language did not predict codability (SI Appendix, S4).

Finally, in their ground-breaking work Brown and Lenneberg (34) identified the most codable stimuli as those with the highest inter-speaker agreement, and also the shortest descriptors. In our study, more codable stimuli within each language received shorter descriptions on average ($r = -0.18$, GAM model $p < 0.001$, see SI Appendix, S4), though there was a “sweet spot”: very short responses were associated with less codable stimuli. This may reflect a balance between a cognitive bias for efficient communication (short labels), on the one hand, and a bias for informativeness (distinct labels), on the other (48, 49).

As with the other analyses, there were significant differences in the strength of this effect between languages and domains, perhaps reflecting cross-linguistic differences in the structure of words or in modality.

To summarize, while vision and sound may be privileged in English, the hierarchy of the senses, as revealed when sampling the diversity of the world's languages, is clearly not the Aristotelian one.

What determines the variation in codability?

If there is not a universal hierarchy of the senses, then what determines the variation found? A list of a-priori hypotheses was compiled about external factors, both demographic and cultural, that could influence codability in each domain. If, as Howes (50) eloquently asserts, the sensorium is "the most fundamental domain of cultural expression, the medium through which all values and practices of a society are enacted", then we ought to see a relationship between specific cultural practices and codability in language (see also 26). To test this, a targeted ethnographic questionnaire was constructed focusing on cultural practices that might predict codability, focusing on each sensory modality separately. Questions included, for example: Does the community use traditional paints or dyes? (predicted to affect color naming; (2)). Is there instruction/training for musical participation (relevant for communicating about sounds)? Do members of the society make pottery, and if so is it patterned? (See SI Appendix, S2 for full set of questions.)In addition, macro-features such as population size, environment, etc. were also tested.

To test the effects of these parameters, they were added as fixed effects into the mixed effects model described above (SI Appendix, S6). We first tested for global effects of macro-demographic features and overall codability across perceptual domains, and found that speakers of languages with a greater number of speakers (estimates from source (53)) had higher agreement across all domains (for a population of 100 estimated mean codability M = 0.14, for a population of 1 million M = 0.24; p = .03), a non-self-evident result. There was a marginal effect of the level of formal education available in the community too: codability was higher with more formal education (high formal schooling availability M = 0.48, low availability M = 0.29; p = .057). However, there was little evidence of a close tie between overall codability and macro-variables such as mode of subsistence, ecology, or environment.

We next tested for specific associations between cultural and macro-features with codability of each perceptual domain. As with the macro-features, only a handful of cultural practices showed reliable associations with codability (Figure 4). In particular, the codability of shape stimuli was higher in communities that make patterned pottery (make patterned pottery M = 0.51, do not make patterned pottery M = 0.2; χ^2 = 9.61, p = .002), and have higher levels of formal education (high education M = 0.51, low education M = 0.19; χ^2 = 6.78, p = .03). In addition, communities that live in square or rectangular houses have better codability for angular shapes than communities living in round houses (angular houses M = 0.35; round houses M = 0.01; χ^2=3.93, p = 0.04).

Sound received higher codability in communities with specialist musicians (specialist musicians M = 0.27, no specialist musicians M = 0.11; χ^2 = 4.10, p = .04). One final predictor of a linguistic codability from cultural parameters appeared in the domain of smell, where subsistence type was a significant predictor (χ^2 = 23.7; p < 0.001): hunter-gatherers had higher codability (M = 0.31) than non-hunter-gatherers (M = 0.10), consistent with previous studies investigating smell in hunter-gatherer societies (30, 54)

Types of responses

In addition to examining agreement in naming, we can ask whether certain domains are more likely to have dedicated lexical resources; another measure of ineffability (55). For example, the hunter-gatherer Jahai predominantly name both colors and odors with domain-specific abstract terminology, whereas English speakers in the same paradigm use basic color words, but ad-hoc source-based descriptions for smells (30). So we examined the sorts of strategies speakers of each language used across perceptual domains.

Once researchers had established the main contentful response for each stimulus, they coded whether each individual participant's response was: Abstract, i.e., a descriptive response that captures the domain-property (e.g., color: *red, green, blue*; smell: *musty, fragrant*; texture: *smooth, rough*); Source-based, i.e., referred to a specific object/source (e.g., color: *gold, silver, ash*; smell: *vanilla, banana*; texture: *fur, silk, beads*; or Evaluative, i.e., gives a subjective response to the stimulus (e.g., *nice, horrible, lovely, yummy*). (See SI Appendix, S1.) To test differences between description types across domains, a Monte Carlo Markov Chain generalized multinomial linear mixed model was used on first responses only (56), predicting type of response by domain with random intercepts for language, stimulus, and respondent (SI Appendix, S7).

Across the board, abstract descriptions were more likely to be used than source-based descriptions (mean percentage of types within each language: abstract $M = 71\%$, source-based $M = 27\%$; $p < .001$), or evaluative descriptions (evaluative $M = 3\%$; $p < .001$; see Figure 5). A mixed effects model testing whether signed languages used a distinct type of response strategy showed no overall effect of the modality of the language; i.e., sign languages were not a distinct group (see SI Appendix, S7). As expected, if a group of speakers were more likely to use abstract terms to refer to a domain, then the codability of that domain was also higher ($r = 0.34$, $\chi^2 = 15.0$, $p < 0.001$, see Figure 4A).

Compared to color, other domains were relatively more likely to elicit evaluative descriptions (all $p < .005$), except for shapes which elicited proportionately more source-based descriptions (all $p < .01$). Smell was significantly more likely to elicit evaluative descriptions than other domains ($p < .005$), consistent with the idea that odor is predominantly distinguished along hedonic lines (57, 58). Only sound was more likely to elicit abstract descriptions than color ($p < .005$); but it notably departed from color and other perceptual domains in predominantly recruiting metaphor for expression.

It is oft-stated that all languages use a *high-low* metaphor to describe variation in pitch (59–62), and this ubiquity of linguistic encoding reflects the fine-tuning of ear anatomy to the environmental statistics of auditory scenes (62). In our sample of diverse languages, however, the most prevalent way to talk about variation in pitch was through the equivalent of a *big-small* metaphor instead (7 languages), followed by *high-low* (4 languages) and *thin-thick* (4 languages). Variations in loudness also primarily elicited a *big-small* metaphor, followed by pairs of non-antonymic contrasts: e.g., *loud-soft, sharp-soft, strong-soft, strong-small*. This suggests that the most "natural" mapping for sound contrasts may, in fact, reside in size rather than spatial location.

Discussion

We conclude that the faculty of language does not constrain, due to intrinsic cognitive architecture, the degree to which different sensory domains are richly coded. Instead, the patterns we found suggest that the mapping of language onto senses is culturally relative. For each perceptual modality there are communities which excel at linguistic expression and those that seem to struggle to put them into words (see Figure 2): American Sign Language and English speakers showed high codability for colors, but Kata Kolok signers and Yélî Dnye speakers struggled, using varied ad-hoc source descriptions; Umpila and British Sign Language participants struggled to name tastes, whereas Farsi and Lao speakers were in total agreement with each other in how to name each tastant. Note also that the modality of language did not predict codability either: while American Sign Language and British Sign

Language looked alike in some ways, Kata Kolok—a village sign language—showed distinct linguistic coding of the senses.

A caveat to this rampant variation is the almost uniformly poor coding of smell across communities (Figure 2 and Figure 3), and its heavy reliance on source-based descriptors (Figure 5). This could reflect the posited "weak link" between smell and language (27, 28), which has led scholars to call olfaction the "muted sense" (63). Main effects such as these are difficult to interpret, however, since low codability could be put down to poor selection of stimuli; although counter to this possibility, these odor stimuli have been used reliably in cross-cultural studies in the past (64). More generally, other studies have shown that odors can be as codable as colors, in particular for hunter-gatherers (30, 54). In this respect, it is striking to see that the hunter-gatherer Umpila from Australia also demonstrated higher codability for smells than colors, suggesting even within the domain of olfaction, there is significant cultural variation.

Overall, we found codability was higher for larger populations, but this association is not straightforward to interpret. It is tempting to suggest that larger communities with a higher likelihood for meeting strangers have a greater need for more specific vocabulary, while small communities can rely more on common ground and shared personal history. However, in our study, population size represents the number of speakers a language has globally, not the size of the community that was sampled. While the two measures are the same for small communities like Umpila or Kata Kolok, they are very different for global languages like English. The variable may instead be a proxy for level of industrialization. Indeed, population size increases significantly with complexity of subsistence type in our sample (Kruskal-Wallis $\chi^2(4) = 12.0$, $p = 0.02$).

Population size may also be a proxy for more centralized states: Political centralization is more likely with larger populations who have higher rates of literacy, standardized education resources, and wider access to canonical culture. In our study, we found evidence consistent with a link between codability and availability of formal education. On the one hand, explicit instruction could have a direct impact on speakers' lexicons, especially for color and shape which are taught in classrooms (65). In fact, we do see a specific effect of education on shape (but not color) terminology. Since our stimulus-set focused on geometric shapes, and these have technical names, perhaps this is unsurprising. On the other hand, the overall pattern suggests a global association between codability scores across perceptual domains and population size, which is harder to explain through formal education, per se, because touch, taste, and smell are typically neglected in the classroom (66).

The exploratory analyses found a few significant links between codability and specific cultural practices for specific domains, but some posited links such as those between color technologies and color codability (6) did not emerge reliably. This does not necessarily demonstrate that codability is independent from cultural influence. Given the diversity of the world's languages, there were relatively few speech communities in our sample. Even though we attempted to sample broadly (e.g., from hunter-gatherer to post-industrial societies), there were few examples of each type, making it difficult to conclude definitively what the specific role of cultural practices might be in the linguistic expression of the senses. For example, there was only one hunter-gatherer community and only one community which lived in round houses. Several variables were also highly clustered: for instance, all communities that had patterned pottery also had leatherwear. Our statistical method was designed to deal with exactly these sorts of facts (see SI Appendix), but the ideal dataset would have more variation in the combination of cultural traits in order to isolate the effect of a particular trait. Alternatively, focused studies comparing closely related communities that differ along one critical dimension could help elucidate the specific relationships between environment, culture, and language (54).

Overall, our study makes clear that there is far more diversity in the linguistic coding of the senses than earlier literature in philosophy and the cognitive sciences had imagined. This is surprising since the intrinsic structure of perceptual pathways, their cumulative phylogenetic history, and the sheer amount of neural tissue dedicated to each might have been expected to heavily imprint accessibility to consciousness and thus the nature of the corresponding linguistic coding. Instead, there is neither a fixed hierarchy of the senses, nor a uniform bifurcation between well-coded distal senses (vision, audition) and the more ineffable more proximal senses (olfaction, and the haptic and gustatory senses). Rather, either by cultural tradition or by ecological adaptation, each language has come to concentrate its efforts on particular sensory domains.

Methods

Sample

We collected data from 20 geographically, typologically, and genetically diverse languages, shown in Figure 1; Table 1. The data come from both small-scale communities, as well as large urban populations. The communities are diverse in their mode of subsistence, including nomadic hunter-gatherers and pastoralists, as well as industrial and post-industrialist societies.

Materials

To assess codability for each perceptual modality a standardized set of materials was constructed according to established psychophysical dimensions (32). There were two different tasks for the visual modality. For color, 80 chromatic Munsell color chips were selected; these were 20 equally spaced hues at 4 degrees of brightness at maximum saturation. On a separate occasion, focal colors were also elicited, using a focal color card. This was a single card with small circles of the same 80 chips used in the free naming laid out according to hue and brightness, plus 4 achromatic chips (67). Participants were also screened for color blindness using Ishihara plates. For shape, there were 20 black and white stimuli presented in a booklet, which included circles, squares and triangles: “good” forms according to Gestalt principles; as well as forms that would not constitute good forms, such as a shape that resembled a flower. Shape stimuli were presented in 2D and 3D variants. Some pages included more than one exemplar. The auditory stimuli consisted of 20 audio files that varied in perceived loudness, pitch and tempo. The stimuli were corrected for perceived loudness—that is, the rising tones had the same sone values, and the loudness scale was corrected to make pitch constant. In the tactile modality, we focused on surface touch, specifically tactile texture. There were 10 texture materials, pressed to a booklet, including materials such as felt, sandpaper, rubber and plastic. There were 5 stimuli for taste, each targeting a “basic” taste: 10 grams of sucrose (sweet), 7.5 grams of sodium chloride (salty), 0.05 grams of quinine hydrochloride (bitter), 5 grams of citric acid monohydrate (sour), and glutamate (umami). Each tastant was dissolved in 100 ml water, except for umami which was presented in powder form. Finally, for smell we used the “The Brief Smell Identification Test™”, a booklet with scratch-and-sniff common odorants, devised for cross-cultural use (68). The test itself is designed to be administered using a forced-choice format but we were interested in people’s free responses and so all text was covered using white tape.

Procedure

Researchers were provided with a manual describing the background, and providing the instructions for running each task (32). Researchers translated the instructions into the target language so that the experiment was conducted in the speaker's native language. For each perceptual modality, participants were asked the equivalent of *What color/shape/sound/etc is this?* In languages where there was not a superordinate term available (equivalent to *color*, for example) other formulations were used, such as *How has it been dyed? How does it strike the eye?* The final questions were always the default way of eliciting the target domain within the language. Researchers were instructed to audio-/video-record sessions for later transcription and coding.

For each domain, the stimulus materials were presented in a single random order across populations, to minimize the effects of order across languages. Researchers were instructed to run the tasks in the following order: color, shape, sound, touch, smell, and taste. On occasion it was not possible to run the experiment in a single sitting and so the experiment was divided into separate sessions. Comparable sound descriptions for auditory stimuli could not be collected for the three sign languages. There is also no data for sound from Yurakaré due to technical problems in the field. It was also not possible to elicit taste descriptions in Mian and Semai as participants did not consent to imbibe the tastants (primarily due to fear of witchcraft).

Coding

Researchers transcribed the data they had collected into the established orthography of the language. For Farsi and Cantonese, a standardized roman orthography was used for ease of comparison. The three sign languages were glossed into English following the usual conventions. Participants' full responses were transcribed, and then for each stimulus the main contentful responses were coded, as well as any modifiers or hedges, using a standardized coding protocol (SI Appendix, S1). So, for example, for the full response *light* *moss green* the main contentful response was coded as *green*, and *light* and *moss* were coded as modifiers. The same procedure was used across modalities.

Acknowledgments.

This work was funded by the Max-Planck-Gesellschaft. SR was supported by a Leverhulme early career fellowship ECF-2016-435; American Sign Language data collection was supported by NIH Grant R01 DC010997; British Sign Language data collection was supported by the Economic and Social Research Council of Great Britain RES-620- 28-6001; RES-620-28-0002: Deafness Cognition and Language Research Centre with the assistance of Adam Schembri, Breish Rowe, Neil Fox, and Ye Tian. For advice with sound stimuli we thank Holger Mitterer, and for color Debi Roberson. Thanks also to Erica Renckens, Nan van de Meerendonk, Nikki de Jong, and others for assistance with stimuli preparation. Data, processing code, and analysis scripts are available in an on-line GitHub repository: https://github.com/seannyD/LoP_Codability_Public.

Author contributions

AM and SCL conceived of the study, and AM, GS and SCL designed the study; AM, LC, GS sourced stimulus materials; AM designed coding protocol in consultation with language researchers; language data was collected and coded as indicated in Figure 1; AM and LC validated coding; SR and AM analyzed data, AM, SR and SCL wrote the paper. The remaining authors are ordered alphabetically by language name for which they contributed data.

References

1. Berlin B, Kay P (1969) *Basic color terms: Their universality and evolution* (University of California Press, Berkeley).

2. Kay P, Regier T (2003) Resolving the question of color naming universals. *Proc Natl Acad Sci U S A* 100(15):9085–9089.

3. Regier T, Kay P, Khetarpal N (2007) Color naming reflects optimal partitions of color space. *Proc Natl Acad Sci* 104(4):1436–1441.

4. Regier T, Kay P, Cook RS (2005) Focal colors are universal after all. *Proc Natl Acad Sci U S A* 102(23):8386–8391.

5. Abbott JT, Griffiths TL, Regier T (2016) Focal colors across languages are representative members of color categories. *Proc Natl Acad Sci* 113(40):11178–11183.

6. Gibson E, et al. (2017) Color naming across languages reflects color use. *Proc Natl Acad Sci*:201619666.

7. Winawer J, et al. (2007) Russian blues reveal effects of language on color discrimination. *Proc Natl Acad Sci* 104(19):7780.

8. Drivonikou GV, et al. (2007) Further evidence that Whorfian effects are stronger in the right visual field than the left. *Proc Natl Acad Sci* 104(3):1097–1102.

9. Franklin A, et al. (2008) Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proc Natl Acad Sci* 105(9):3221–3225.

10. Roberson D, Pak H, Hanley JR (2008) Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition* 107(2):752–762.

11. Thierry G, Athanasopoulos P, Wiggett A, Dering B, Kuipers J-R (2009) Unconscious effects of language-specific terminology on preattentive color perception. *Proc Natl Acad Sci* 106(11):4567–4570.

12. Siok WT, et al. (2009) Language regions of brain are operative in color perception. *Proc Natl Acad Sci* 106(20):8140–8145.

13. Clark WELG (1959) *The antecdents of man* (Edinburgh University Press, Edinburgh).

14. Cartmill M (1974) Rethinking primate origins. *Science* 184(4135):436–443.

15. Barton RA, Purvis A, Harvey PH (1995) Evolutionary radiation of visual and olfactory brain systems in primates, bats and insectivores. *Philos Trans Biol Sci* 348(1326):381–392.

16. Dominy NJ, Ross CF, Smith TD (2004) Evolution of the special senses in primates: Past, present, and future. *Anat Rec* 281A(1):1078–1082.

17. Denion E, Hitier M, Guyader V, Dugué A-E, Mouriaux F (2015) Unique human orbital morphology compared with that of apes. *Sci Rep* 5(1). doi:10.1038/srep11528.

18. McGann JP (2017) Poor human olfaction is a 19th-century myth. *Science* 356(6338):eaam7263.

19. Gilad Y, Wiebe V, Przeworski M, Lancet D, Pääbo S (2004) Loss of olfactory receptor genes coincides with the acquisition of full trichromatic vision in primates. *PLoS Biol* 2(1):e5.

20. Viberg Å (1984) The verbs of perception: A typological study. *Explanations for Language Universals*, eds Butterworth B, Comrie B, Dahl Ö (Mouton de Gruyter), pp 123–162.

21. San Roque L, et al. (2015) Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cogn Linguist* 26:31–60.

22. Howes D (2005) *Empire of the senses* (Berg, Oxford; New York).

23. Kant I (2006) *Anthropology from a pragmatic point of view* (Cambridge University Press, Cambridge).

24. Sperber D (1975) *Rethinking Symbolism* (Cambridge University Press, Cambridge). English translation by Alice L. Morton.

25. Henning H (1916) *Der Geruch* (Johann Ambrosius Barth, Leipzig).

26. Reid T (1764) *An inquiry into the human mind on the principles of common sense* (Thomas Tegg, London).

27. Olofsson JK, Gottfried JA (2015) The muted sense: Neurocognitive limitations of olfactory language. *Trends Cogn Sci* 19(6):314–321.

28. Yeshurun Y, Sobel N (2010) An odor is not worth a thousand words: From multidimensional odors to unidimensional odor objects. *Annu Rev Psychol* 61:219–241.

29. Classen C, Howes D, Synnott A (1994) *Aroma: The cultural history of smell* (Routledge, London).

30. Majid A, Burenhult N (2014) Odors are expressible in language, as long as you speak the right language. *Cognition* 130(2):266–270.

31. Kay P, Berlin B, Maffi L, Merrifield WR, Cook R (2009) *The world color survey* (CSLI Publications, Stanford, CA).

32. Majid A ed. (2007) *Field Manual Volume 10* (Max Planck Institute for Psycholinguistics, Nijmegen).

33. Brown P (2011) Color Me Bitter: Crossmodal Compounding in Tzeltal Perception Words. *Senses Soc* 6(1):106–116.

34. Brown RW, Lenneberg EH (1954) A study in language and cognition. *J Abnorm Soc Psychol* 49(3):454–462.

35. Agrillo C, Roberson D (2009) Colour language and colour cognition: Brown and Lenneberg revisited. *Vis Cogn* 17(3):412–430.

36. Lucy JA (1992) *Language diversity and thought: A reformulation of the linguistic relativity hypothesis* (Cambridge University Press, Cambridge).

37. Simpson EH (1949) Measurement of diversity. *Nature* 163:688–688.

38. Majid A, Gullberg M, van Staden M, Bowerman M (2007) How similar are semantic categories in closely related languages? A comparison of cutting and breaking in four Germanic languages. *Cogn Linguist* 18(2):179–194.

39. Shannon CE (1949) Communication theory of secrecy systems. *Bell Labs Tech J* 28(4):656–715.

40. Bates D, Maechler M, Bolker B, Walker S (2014) lme4: Linear mixed-effects models using Eigen and S4. *R Package Version* 1(7). Available at: http://keziamanlove.com/wp-content/uploads/2015/04/StatsInRTutorial.pdf [Accessed March 11, 2016].

41. R Core Team (2013) *R: A language and environment for statistical computing* (R Foundation for Statistical Computing, Vienna, Austria) Available at: http://www.R-project.org.

42. Henning H (1916) The quality series of taste. *Z Fuer Psychol* 8(1):203–219.

43. Clark A (1993) *Sensory qualities* (Clarendon Press, Oxford).

44. Stevens SS, Davis H (1938) *Hearing: Its psychology and physiology* (Wiley, Oxford).

45. Keller A (2016) *The philosophy of olfactory perception* (Palgrave Macmillan, Cham).

46. Frasnelli J, Collignon O, Voss P, Lepore F (2011) Crossmodal plasticity in sensory loss. *Progress in Brain Research* (Elsevier), pp 233–249.

47. Sela RJ, Simonoff JS (2011) *REEMtree: Regression trees with random effects. R package version 0.90.3.*

48. Regier T, Kemp C, Kay P (2015) Word meanings across languages support efficient communication. *The Handbook of Language Emergence*, eds MacWhinney B, O'Grady W (Wiley, Hoboken, NJ), pp 237–264.

49. Zipf GK (1949) *Human behavior and the principle of least effort* (Addison-Wesley, Cambridge).

50. Howes D (2003) *Sensual relations: engaging the senses in culture and social theory* (University of Michigan Press).

51. Regier T, Carstensen AB, Kemp C (2016) Languages support efficient communication about the environment: Words for snow revisited. *PLoS One* 11(4):e0151138.

52. Levinson SC (2000) Yélî Dnye and the theory of basic color terms. *J Linguist Anthropol* 10(1):3–55.

53. Simons GF, Fennig CD eds. (2017) *Ethnologue: Languages of the world* (SIL International, Dallas, Texas). Twentieth Available at: http://www.ethnologue.com.

54. Majid A, Kruspe N (2018) Hunter-gatherer olfaction is special. *Curr Biol*.

55. Levinson SC, Majid A (2014) Differential ineffability and the senses. *Mind Lang* 29(4):407–427.

56. Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: The mcmcglmm R package. *J Stat Softw* 33(2):1–22.

57. Khan RM, et al. (2007) Predicting odor pleasantness from odorant structure: Pleasantness as a reflection of the physical world. *J Neurosci* 27(37):10015–10023.

58. Winter B (2016) Taste and smell words form an affectively loaded and emotionally flexible part of the English lexicon. *Lang Cogn Neurosci* 31(8):975–988.

59. Stumpf C (1883) *Tonpsychologie* (Hirzel, Leipzig).

60. Pratt CC (1930) The spatial character of high and low tones. *J Exp Psychol* 13(3):278–285.

61. Evans KK, Treisman A (2011) Natural cross-modal mappings between visual and auditory features. *J Vis* 10(1):6–6.

62. Parise CV, Knorre K, Ernst MO (2014) Natural auditory scene statistics shapes human spatial hearing. *Proc Natl Acad Sci* 111(16):6104–6108.

63. Ackerman D (1990) *A natural history of the senses* (Vintage Books, London).

64. Doty RL, Applebaum S, Zusho H, Settle RG (1985) Sex differences in odor identification ability: A cross-cultural analysis. *Neuropsychologia* 23(5):667–672.

65. de Sousa H (2011) Changes in the Language of Perception in Cantonese. *Senses Soc* 6(1):38–47.

66. Majid A, Manko P, de Valk JM (2017) Language of the senses. *Scientific Breakthroughs in the Classroom!*, ed Dekker S (Science Education Hub Radboud University, Nijmegen), pp 40–76.

67. Majid A (2008) Focal colours. *Field Manual Volume 11*, ed Majid A (Max Planck Institute for Psycholinguistics, Nijmegen), pp 8–10.

68. Doty RL (2001) *The Brief Smell Identification Test administration manual* (Sensonic Inc, White Horse Pike, NJ).

Figure Legends

Figure 1: Languages (and researchers) contributing to the study. Locations indicate fieldsites where data was collected.

Figure 2: Boxplots of codability (measured by Simpson's diversity index) plotted by domain and language (0 indicates low codability; 1 indicates high codability). English shows the predicted high codability for color, shape and sound, and low codability for touch taste and smell; but other languages exhibit different hierarchies.

Figure 3: The hierarchy of the senses across languages according to the mean codability of each domain, with the presumed universal Aristotelian hierarchy on top. There is no universal hierarchy of the senses across diverse languages worldwide.

Figure 4: Factors that explain codability: (A) Codability is higher for a domain if more abstract terms are used to refer to it (regression line from a mixed effects model). (B) Codability is higher for larger populations (raw data with regression line from a mixed effects model), and (C) communities with formal schooling. (D) Codability of sounds is higher for communities with specialist musicians. (E) Codability for shape is higher for communities with more access to formal schooling and (F) patterned pottery. (G) Codability for angular shapes is higher for communities that live in angular houses. (H) Hunter-gatherers have higher codability for smell than other communities.

Figure 5: Strategies for describing perceptual stimuli across languages. For each domain and language, the proportion of Abstract, Source-based, and Evaluative responses are plotted.