Carving the body at its joints: Does the way we speak about the body shape the way we think about it?*

SIMON DEVYLDER [iD]

Lund University and UiT the Arctic University of Norway

CHRISTOPH BRACKS [iD]

University of Cologne

MISUZU SHIMOTORI

University of Gothenburg and University of Bergen

A N D

POPPY SIAHAAN

University of Cologne

ABSTRACT

Looking at the way different linguistic communities speak about a universally shared domain of experience raises questions that are central to the language sciences. How can we compare meaning across languages? What is the interaction between language, thought, and perception? Does linguistic diversity entail linguistic relativism? The literature on the naming systems of the body across languages have addressed these questions with little consensus. In the present study, we contribute to this debate with a comparison of body part terms in French, Indonesian, and

577

Japanese. Using an updated version of the body coloring task, we observed both diversity and cross-linguistically shared patterns. Importantly, we also observed that speakers of languages which violate the wrist/ankle joint boundary rule do not collapse the distinction in thought. This key finding goes against the conflation of language and thought and leads us to conclude that linguistic diversity does not entail linguistic relativism. Methodologically, we advocate for the use of a culturally neutral etic space as a necessary tool in semantic typology. Theoretically, we propose that language is a multilevel phenomenon, which results from the interaction of non-linguistic and cross-culturally shared embodied motivations, context-specific situated language use, and culturally specific sedimented linguistic conventions.

KEYWORDS: linguistic diversity, language and thought, categorization, semantic typology, body parts

1. Introduction

Concepts must carve nature at its joints, Socrates preached (Plato, 265e). Does the maxim apply for the categorization of the body in perception, thought, and language? The sensorial reality of perceptual boundaries in the segmentation of the body at its joints has been evidenced by studies in cognitive neuroscience (Shen, Smyk, Meltzoff, & Marshall, 2018) and in behavioral psychology (e.g., de Vignemont, Majid, Jola, & Haggard, 2009; Knight, Cowie, & Bremner, 2017; Knight, Longo, & Bremner, 2014), which show that joints, such as the wrist, have a category boundary effect on the perception of the body. Brown (1976) and Andersen (1978) notoriously claimed that the segmentation of the body lexicon is motivated by universal principles like shape, size, or visual discontinuities, and that body part terms are the linguistic labels of the perceptually predetermined conceptual parts. 'From perception to language via conceptual thought' sounds like a path upon which many psychologists, linguists, and anthropologists would be ready to embark. That path would naturally take us to the following syllogism: the same type of perceptual system is shared by all human beings, semantic boundaries come from perceptual and conceptual boundaries, therefore all languages segment the body according to the same universal rules. But, do they?

The contributors to a Special Issue of *Language Sciences* (Enfield, Majid, & van Staden, 2006) provide evidence against that intuition. The series of studies challenged the universalist claims by providing evidence for the existence of a great diversity of linguistic categorization of the body. These findings and those of a later publication (Majid & van Staden, 2015) are summarized in Table 1.

This body of literature concludes that there are a number of lexical gaps found in a variety of languages, and thereby a great diversity of body part naming systems across languages that goes against the universal assumptions of Brown (1976) and Andersen (1978). Wierzbicka (2007) disagrees. According to

TABLE 1. *Variation of the semantic extensions of body part terms across eight languages (adapted from Majid & van Staden, 2015)*

	Upper Arm	Lower Arm	Hand	Upper Leg	Lower Leg	Foot
Jahai (Burenhult, 2006)	*bliŋ*	*ɟrbɛr*	*cyas*	*blɨʔ*	*gor*	*can*
Punjabi (Majid, 2006)	*baa*		*hatth*	*lətt*		*pær*
Dutch (Majid & van Staden, 2015)	*arm*		*hand*	*been*		*voet*
Japanese (Majid & van Staden, 2015)	*ude*		*te*	*ashi*		
Yélî Dnye (Levinson, 2006)	*kêê*			*kpââli*	*yi*	
Indonesian (Majid & van Staden, 2015)	*tangan*			*kaki*		
Savosavo (Wegener, 2006)	*kakau*			*nato*		
Lavukaleve (Terrill, 2006)	*tau*					*fe*

Wierzbicka, a lexical gap for 'hand' would mean a conceptual gap, which is impossible due to the experiential salience of this body part. Consequently, she explains that if a language only has one word to refer to both *concepts* 'hand' and 'arm' (e.g., *ręce* in Polish) it is because this single word has a plurality of meanings, i.e., is polysemous. Hence the language diversity advocates are wrong in defending semantic generality, and are also guilty of exoticizing these non-Western world languages by suggesting their speakers have no concept for the words that are missing in their body lexicon. Levinson (2006), a contributor to the Special Issue, did however write that the lexical gaps are not conceptual gaps. But in Wierzbicka's defense, contributors to the Special Issue did not make a systematic distinction between language and conceptualization: "[w]e offer this collection as a step in reviving interest in the empirical study of the way in which human beings C O N C E P T U A L I Z E and C A T E G O R I Z E their bodies as physical entities with parts" [our emphasis] (Enfield et al., 2006, p. 146). It is thus not exactly clear if the findings reported in Table 1 display the great diversity of body categorization in language only, or also in thought. Majid (2015, p. 376) takes a clearer stand on the relation between language and thought and points out that "it may well be that every speaker has a distinct non-linguistic concept of hand, but not all concepts are reflected in the lexicon". As Majid points it out, when Wierzbicka (2007, p. 29) writes that "the most reliable evidence for the presence of such a concept is the presence of a word", this arguably constitutes one of the main weaknesses of the universalist argument because it conflates (linguistic) meaning and (non-linguistic) concepts. At the core of this debate lies two broad theoretical questions: that of linguistic diversity vs. language universals, and that of the interaction between language and thought.

The methods employed to address these questions are also at the crux of the debate. Wierzbicka harshly criticizes the methods used by the diversity group as reductionist and theoretically empty. According to her, the decontextualized nature of the body coloring method cannot establish the polysemy of a word (Wierzbicka, 2007, p. 17). She claims that a more valid methodological alternative would be that of the natural semantic metalanguage (NSM) framework. The NSM method is based on a "mini-language" (Wierzbicka, 2007, p. 18; Wierzbicka & Goddard, 2017, p. 32) that describes the 65 universal semantic primes of the NSM framework. This is indeed a theoretically heavy method in contrast to the "theoretical vacuum" (Wierzbicka, 2007, p. 17) that the diversity contributors are charged with. The validity of Wierzbicka's critique and the solution the NSM has to offer can, however, be questioned. Arguably, a fundamental prerequisite for comparing meaning across languages (i.e., doing semantic typology) is the use of an E T I C , culturally neutral, and hence non-language-specific space, against which different meanings can be compared. The mini-language used by the NSM, which has as many versions

580

as there are languages, works against any possibility for a neutral etic space. No matter how constrained the use of these 'universal language-specific labels and definitions' are, they inherently run the risk of introducing as many EMIC biases as there are languages. In contrast, the non-linguistic nature of the body coloring task provides a culturally neutral etic space advocated by others (Majid & van Staden, 2015, p. 586).

The present paper aims to move the debate forward by addressing the following general research questions:

- RQ1. How can we compare meaning across languages?
- RQ2. Are body lexicons linguistically diverse?
- RQ3. Does language shape thought?

It would be naively ambitious to promise answers to these fundamental questions within the scope of a single paper. We only aim to address RQ1 by motivating the methodological decisions we made with a set of principles that we think are necessary, albeit not sufficient, to include in methods of semantic typology in general, and in the study of body lexicons across languages in particular (Section 2). RQ2 is addressed by implementing these principles in a study where 90 speakers of three unrelated languages (French, Indonesian, and Japanese) are asked to perform our 'version 2.0' of the body coloring task, which was simultaneously video-recorded. The results of this study, reported in Section 3, demonstrate both diversity and cross-linguistically shared patterns in our limited sample of languages. To help us address RQ3, we also report an analysis of the video-recorded performances in Section 3. Specifically, we provide evidence for a case of misalignment between language and thought: the lexical units that collapse the hand–arm (*tangan*) and leg–foot (*kaki*) distinctions in Indonesian do not match the discontinuous coloring performances of its speakers, who color the upper and lower limb in two clear sequences at the level of the wrist and ankle joints, respectively. This important finding suggests that the language we speak does not determine the way we think about the body and its parts, thereby arguing against the conflation of language and thought. We take stock of these findings in Section 4 and discuss their broader methodological and theoretical implications for semantic typology. We offer some concluding thoughts in Section 5.

2. Methods

The stimuli, procedure, and analysis are developed from the body coloring task originally designed in van Staden and Majid (2006) and applied in Majid and van Staden (2015).

2.1. LANGUAGES AND PARTICIPANTS

Data were collected from 90 adult L1 speakers of French (30), Indonesian (30), and Japanese (30). A criterion for recruiting participants was that they should not be educated in medical studies, as their description would not reflect the knowledge of the human anatomy of the average population. Particular attention was paid to collecting the data at the same place and from speakers who grew up speaking French, Indonesian, and Japanese as their first language. These recruitment criteria are particularly important for the representativeness of the Indonesian data. Indonesian is the national language of Indonesia, and the official language of the country's administration, government, and mass media (Sneddon, Adelaar, Djenar, & Ewing, 2012). There are substantial regional variations within Indonesian in part due to the fact that it is mostly acquired as L2 by speakers whose L1 is a regional language (e.g., Balinese, Batak, Javanese, Sundanese, and the numerous languages of East Indonesia and Papua) (Tadmor, 2017). What is referred to in this paper as Indonesian is more specifically colloquial Jakarta Indonesian (Sneddon, 2006), which is a variety that is spoken in Jakarta. It is the language of television series and movies, and therefore understood throughout Indonesia. We selected participants who grew up and live in Jakarta, speaking colloquial Jakarta Indonesian as their first language. The Indonesian data was thus collected in Jakarta, the French data was collected in Lyon, France, and the Japanese data in Tokyo, Japan.

2.2. STIMULI

The stimuli were compiled in a booklet and consisted of identical body pictorial representations printed on separate sheets of paper. The body representations include both the front and the back of an androgenized and physiognomically neutral human being. A body part term in the target language appeared centered at the top of each sheet (Figure 1).

For the present study, we selected 18 to 20 body part terms per language on the basis of their frequency of use in an elicitation task in which 30 French, Indonesian, and Japanese participants had to describe pictures of injuries located on various areas of the body (Devylder, Kozaï, & Siahaan, unpublished observations). The stimuli are summarized in Figure 2, indicating where the injury was located on the human body with what was depicted to be the cause of the injury in brackets.

By selecting the body part terms to include in this way, we were able to avoid any presupposition on the hierarchical status of these terms. For example, it is unclear why Majid and van Staden (2015) omitted the Indonesian body part term *lengan*, mentioned in Brown (1976, p. 418). The authors may have considered the term to be hierarchically dependent on *tangan*, in the same way that
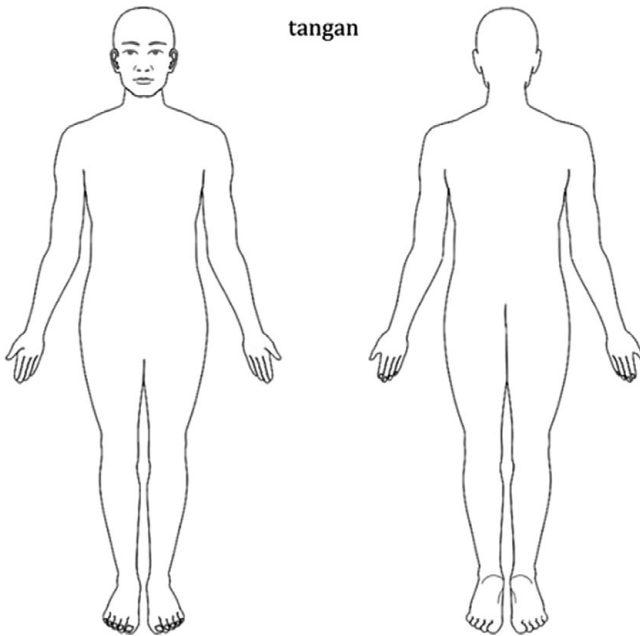
582

Fig. 1: Example of a stimulus from one of the Indonesian booklets.

forearm is part of *arm* in English. However, there is no empirical evidence supporting the existence of such a hierarchical relation between *tangan* and *lengan*: neither dictionary entries (Stevens & Schmidgall-Tellings, 2004, pp. 575, 995)[1] nor L1 Indonesian speaker intuitions suggest that *tangan* is more basic than *lengan*. Given that Wierzbicka's (2007) core argument against the diversity of body categorization in language is anchored in the claim that all languages have a term for 'hand', the semantic extension of *lengan* is both too uncertain and too central to the debate to be omitted from the stimuli a priori.

Moreover, Majid and van Staden (2015, pp. 575–576) only included the Japanese term *ashi* 足 in the design of their task, and excluded its homophone *ashi* 脚. Both terms have the same pronunciation but are written with different characters, which suggests a potentially different meaning. Following the same rationale as above, both characters were included in the Japanese elicitation booklets. In sum, a guiding principle in the design of semantic typology elicitation tools should perhaps be 'when in doubt, leave it in'.

As the original procedure indicates (van Staden & Majid, 2006), participants who have just colored in a term for *foot* from the ankle to the toes could be

---

[1] *Lengan* "1. Arm, upper limb of the human body from the shoulder to the hand" ; *Tangan* "hand, arm"
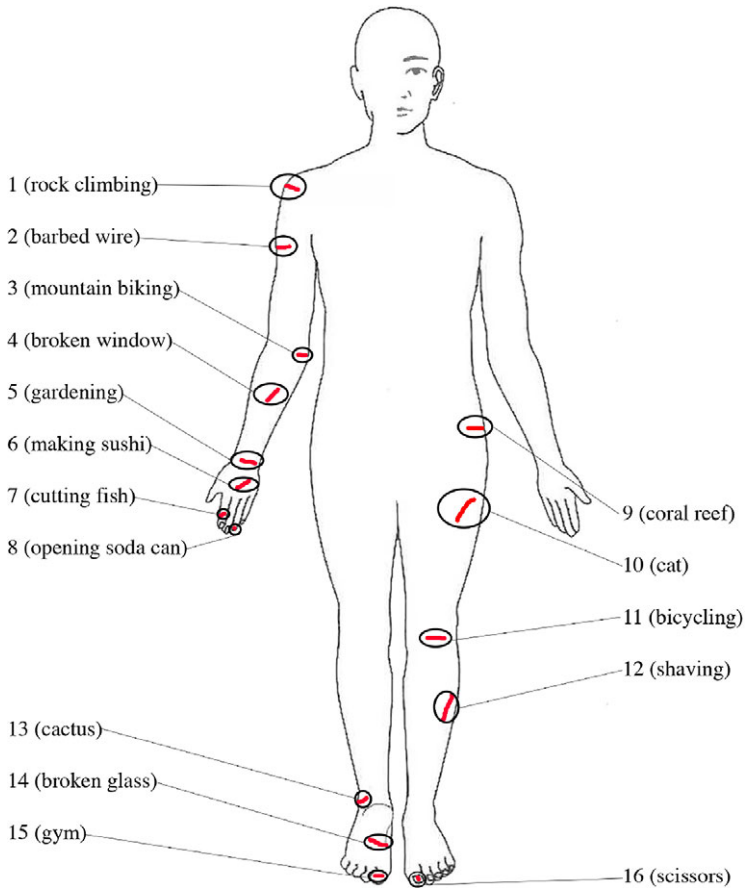
583

Fig. 2: Summary of the elicitation stimuli from Devylder, Kozaï, and Siahaan (unpublished observations).

reluctant to include the ankle in the sheet where they are asked to color in a term for *leg*. We wanted to measure whether the order of presentation affected participants' performances and thus split the 30 participants of each target language into three groups. The first group had a randomly organized booklet; the second group was given a booklet that organized the body part terms from larger to smaller body segments and from the top down (e.g., arm > shoulder > upper arm > elbow > forearm > wrist > hand > finger > fingernail); and the third group was given a booklet that was arranged from smaller to larger body segments and from the bottom up (e.g., fingernail > finger > hand > wrist > forearm > elbow > upper arm > shoulder > arm).

584

2.3. ELICITATION PROCEDURE

Participants were given booklets and instructions to color in the body part that is named on each page. All instructions and communication between the experimenter and the participants were done in the target language. The instructions were to color in all of the body part, and only the body part listed on each page. Participants started with five training pictures, during which the experimenter calibrated the experiment if needed. We replaced the fine-liner blue pen of the 2015 procedure with a wedge-shaped highlighter to allow participants both to color large areas with minimal effort and to precisely draw boundaries with the tip of the wedge shape. All experiments were video-recorded. We thus had two types of raw data to code: the colored-in booklets and the video-recorded performances.

2.4. THE ANATOMICAL GRID AS AN ETIC YARDSTICK

Using English as a 'yardstick' is economical and convenient because we can assume that all readers know what the English terms *hand* and *foot* refer to, more or less. While the use of English is fine and perhaps even necessary to summarize studies, however, it should be avoided both in the analysis and in the interpretation of results. For example, as mentioned in the 'Introduction', the NSM approach is based on a "mini-language" (Wierzbicka, 2007, p. 18; Wierzbicka & Goddard, 2017, p. 32) often presented in its English version, but which has "as many versions as there are languages". The problem of using English as a universal yardstick is already problematic in semantic typology because we do not know the length of the stick, and therefore cannot really measure anything against it. Of course, using as many yardsticks as there are languages to demonstrate universal tendencies across languages makes the validity of cross-linguistic analyses very difficult to achieve. Hence the call from Majid and van Staden (2015, p. 586) to use "a neutral etic space (a non-linguistic representation) for better capturing the otherwise invisible differences in reference between specific languages".

To answer that call, we developed a non-linguistic and culturally neutral system in the form of the anatomical grid illustrated in Figure 3. The goal was to find a way to segment the body into zones that were not bound to any language-specific categorization of the body. We therefore segmented the upper and lower limbs into zones based on the anatomical boundaries formed by the multiplicity of muscles, tendons, and bones that constitute the human body. So, for instance, zone A1 corresponds to the *clavicule + platysma*, zone A2 to the *pectoralis major*, zone A3 to the *deltoid*, and so on. This anatomical grid consisted of 63 zones (the complete description of which is listed in 'Appendix 1') and allowed us to measure the semantic
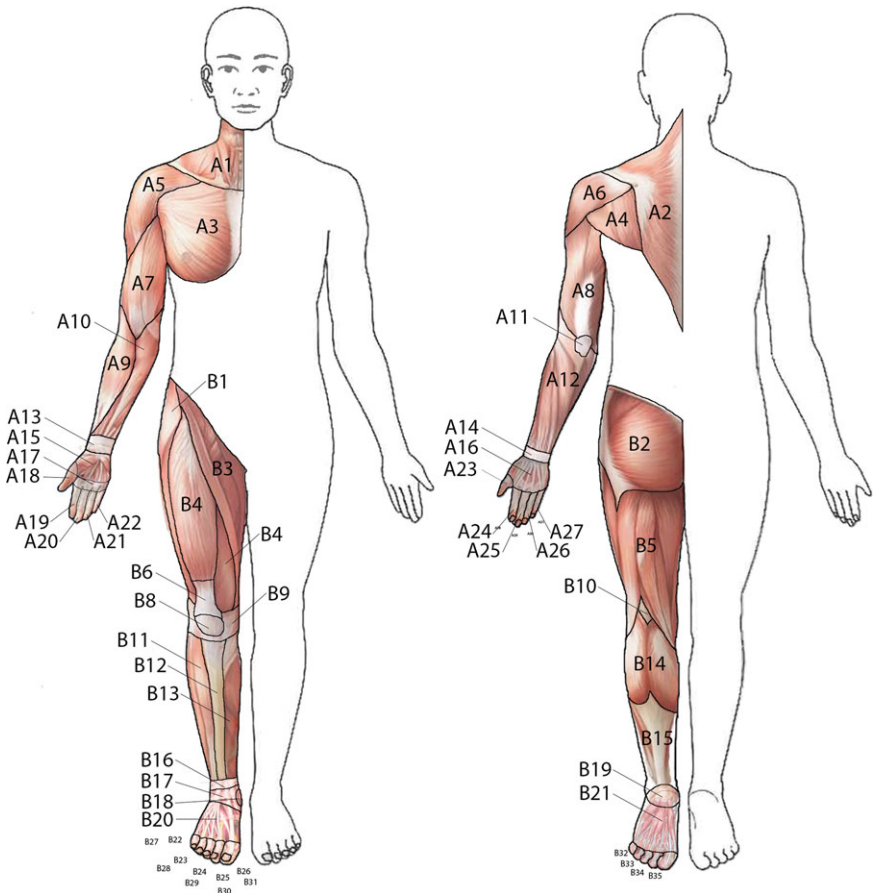
585

Fig. 3: The anatomical grid.

extensions of body part terms and to compare them across languages against a stable and etic yardstick.

2.5. ANNOTATION PROCEDURE OF THE COLORED-IN BOOKLETS

We scanned all colored-in booklets, and superimposed the anatomical grid on each colored-in body schema in Adobe Photoshop. We then measured the surfaces that were colored by each participant for each body part term with the Adobe Photoshop pixel counting tool. This tool allowed us to measure the colored-in surface for each term and for each participant (e.g., Participant FRE-2S-12 colored in 1,024 pixels of A3 for the term *bras*) and the maximum

586

pixel density for each anatomical zone (e.g., A3 = 39,693 pixels). Our raw data thus consisted of three large tables (i.e., one per language) with anatomical zones as columns, individual body part terms colored in per participant as rows, and pixel density as the value for each cell. By distributing the total colored surface over anatomical zones, the anatomical grid allowed us to locate the semantic extension of a body part term more precisely than with the 2015 procedure. In contrast, the 2015 procedure could only compare the total extent of colored-in surfaces, and could not provide quantifiable information on the actual location of the colored-in surface. For example, we could have two identical colored-in surfaces (e.g., c.25,000 pixels) for the Dutch term *arm* and the Japanese term *ude* but no indication about the location of the colored-in surface on the upper limb (e.g., one surface could start from the shoulder and end mid-forearm, and the other start lower but extend further down the wrist region: both would have similar total surfaces but different semantic extensions).

2.6. ANNOTATION PROCEDURE OF THE VIDEO-RECORDED PERFORMANCES

The video-recordings of the body coloring task performances allowed us to capture a phenomenon that could not have been observed by only looking at the final result of the task (i.e., the completed colored-in booklets). The video data was imported into ELAN (Lausberg & Sloetjes, 2009) and annotated as follows. If a participant's coloring performance was done in one continuous fashion, the corresponding elicited body part term was assigned the value 0. If a participant's coloring performance was done in two clearly distinct sequences (i.e., by marking a pause at the level of the wrist joint for more than 1.5 seconds), then the corresponding elicited body part term was assigned the value 1. Example 1 illustrates such discontinuous coloring performances.
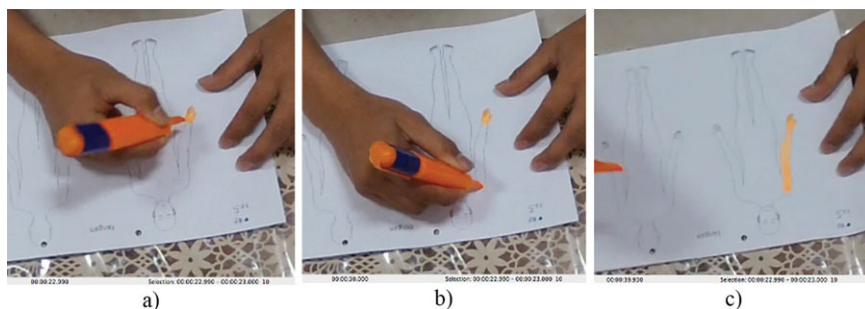
Example 1. Indonesian participant coloring *tangan* in two sequences.

587

In Example 1, the participant starts by coloring the extremity of the upper limb until approximately zone A13 (i.e., the wrist) and lifts up his pen at 00:22:990 seconds (1a). He then marks a long pause of 7 seconds until he starts coloring again at 00:30:000 (1b) from the top of the upper limb and finishes his coloring performance of *tangan* at 00:39:930 (1c) before moving on to the next page of the booklet. This coloring performance of *tangan* was thus given the value 1.

3. Results

We address RQ2 by asking our dataset if there are body part terms that collapse the wrist/ankle boundaries and terms that mark the distinction without being hierarchically dependent of one another in French (FRE), Indonesian (IND), and Japanese (JPN). To address this research question, we limited our analyses to terms that were likely to meet these criteria according to the previous literature (e.g., Brown, 1976; Majid & van Staden, 2015; Wierzbicka, 2007) or which could not be clearly excluded from meeting them (e.g., the existence of *lengan* (IND) and *ashi* 脚 (JPN)). We first report the semantic extensions of this list of body part terms within language groups (3.1) before comparing them across languages (3.2). Finally, we analyze the results of the video-recorded performances (3.3).

3.1. THE SEMANTIC EXTENSIONS OF BODY PART TERMS WITHIN LANGUAGE GROUPS

3.1.1. *French*

As described in Section 2.5, each anatomical zone has a total surface measured in pixel density. For example, A7 (i.e., the biceps muscle) has a total density of 20,303 pixels. As can be seen in Chart 1, the area of A7 colored in by the

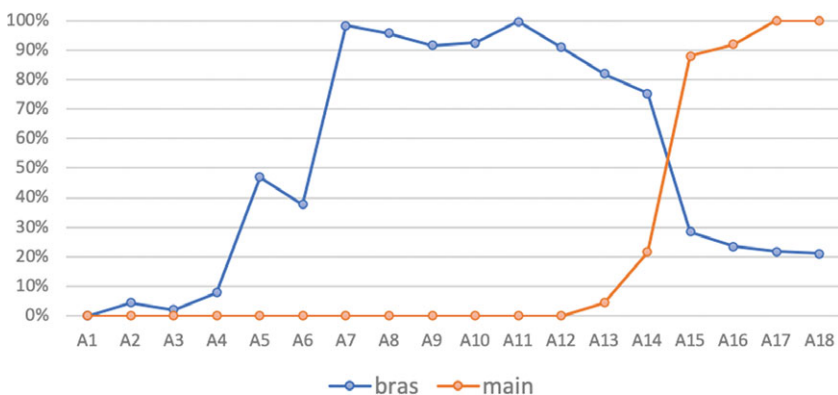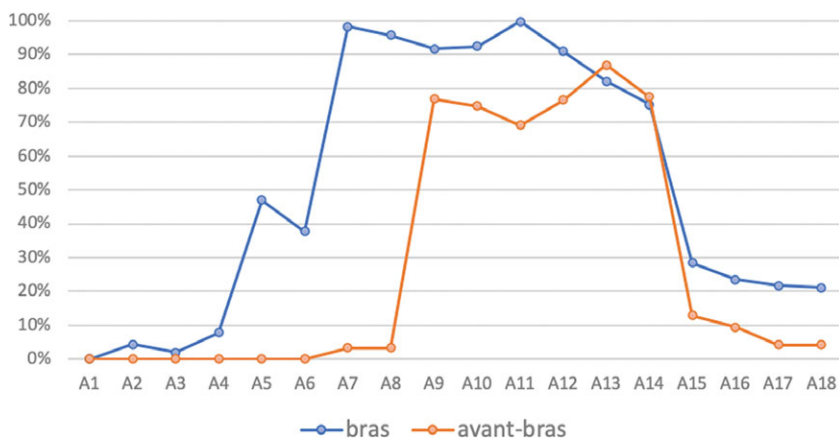Chart 1

588

Chart 2

30 French participants, when asked to indicate the surface of the body corresponding to the term *bras*, measured on average 19,897 pixels, or 98% of the total surface of A7. In contrast, the area of A7 colored in by the same participants, when asked to indicate the surface of the body corresponding to *main*, measured 0 pixel, or 0% of A7.

For *bras*, there was a decrease in pixel density between A14 and A15 that coincided with an increase in pixel density between the same zones for *main*. This pattern suggests that the meanings of *bras* and *main* do not overlap in French. Because they are both contiguous and largely non-overlapping, we can further conclude that the meaning of *bras* and *main* are not hierarchically dependent on one another (e.g., one is not part of the other) and at same level of partonomic depth (i.e., level B in Andersen, 1978; Majid, 2010). In contrast, the comparison of the semantic extensions of *bras* and *avant-bras* in Chart 2 is an example of two semantic extensions that overlap (from A9 to A14), where the colored area of the latter was larger than the colored area of the former. We can interpret this overlap as an index for a hierarchical relation between the two terms (e.g., *avant-bras* 'forearm' is a part of *bras* 'arm').

As described in Section 2.2, each language group was split into three groups who were given a differently ordered booklet. We compared whether the area that the French speakers colored-in for *bras* was affected by the order of presentation of the stimuli by conducting a one-way ANOVA with order of presentation (random, top-down, bottom-up) as the independent variable and the pixel density of the colored anatomical zone A15 as the dependent variable. We selected A15 as the dependent variable because it was the point at which the average proportion of pixel density decreased in the coloring performance of *bras* (Chart 1). In this way, we were able to assess whether French participants
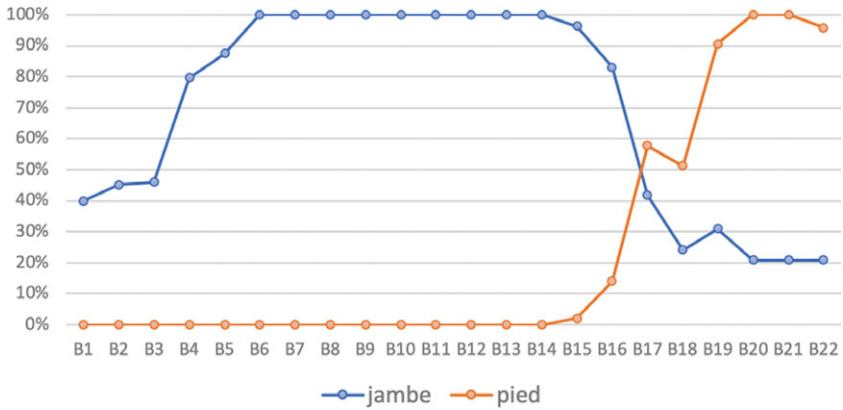
589

Chart 3

who have been asked to color in *main* before *bras* (i.e., the bottom-up group) were more likely to extend the meaning of *bras* to the tip of the fingers than participants who were given a differently ordered booklet. The one-way ANOVA revealed no significant effect of order of presentation on the pixel density of A15 ($F(2,27) = 0.94$, $p > .405$).

As with any study based on a limited sample of a population, the semantic extensions reported in this section are to be taken with a pinch of salt with regard to their representative power. The S H A N N O N β - D I V E R S I T Y I N D E X (Whittaker, 1972) can tell us how big that pinch is. The β-diversity index is obtained by measuring the number of distinct anatomical zones that participants colored-in for each term. Since we have 30 participants in each language group, the β-diversity index ranges from 1.00, indicating the exact same distribution of the number of colored-in anatomical zones for a body part term by all 30 participants (i.e., the semantic extension of the body part term is exactly the same), to 30.00, indicating divergent distributions, without any overlap between the participants' performances (i.e., 30 completely different semantic extensions of the same body part term). In sum, the higher the β-diversity index, the lower the consistency within a language group per elicited body part term, such that $\beta = 30.00$ means maximally inconsistent answers, and $\beta = 1.00$ means maximally consistent answers (see Whittaker, 1972, pp. 232–235) for the formula, and a detailed account of how to calculate this metric). The β-diversity index for the coloring performances was 1.78 for *bras* and 1.02 for *main*. Taken together, the lack of effect of presentation order and the β-diversity index suggest that the semantic extensions of *bras* and *main* are quite robust and not subject to significant intrapersonal variation in French.

Chart 3 reveals the same pattern observed in the comparison of *main* and *bras* for the lower limb terms *jambe* and *pied*. The pixel densities decreased and

590

increased with the same magnitude at the same point and largely did not overlap. We thus conclude that *jambe* and *main* are not hierarchically dependent on one another and are at same level of partonomic depth.

A one-way ANOVA revealed no significant effect of order of stimulus presentation (random, top-down, bottom-up) on the pixel density of B20 ($F(2,27) = 0.59$, $p > .563$). We selected B20 as the dependent variable because it was the zone where the average proportion of pixel density decreased in the coloring performance of *jambe* (Chart 3). The ANOVA thereby indicates that the presentation order had no effect on excluding the '*pied* segment' from the semantic extension of *jambe*. The β-diversity index for the coloring performances was 1.54 for *jambe* and 1.10 for *pied*. We can conclude that the semantic extensions of *jambe* and *pied* are quite robust in French.

In sum, French has two distinct terms for the upper limb and two distinct terms for the lower limb that do not stand in a hierarchical relation with each other and that semantically mark the wrist joint boundary and the ankle joint boundary, respectively.

3.1.2. *Japanese*

The semantic extensions of the body part terms *ude* 腕 and *te* 手, as elicited by the body coloring task, are reported in Chart 4. If we compare the average proportions of colored-in anatomical zones for *ude* 腕 and *te* 手, we notice a similar pattern to the French coloring for the two upper limb terms. There was a clear decrease in pixel density between A14 and A15 for *ude* 腕 that coincided with an increase in pixel density between the same zones for *te* 手. We thus
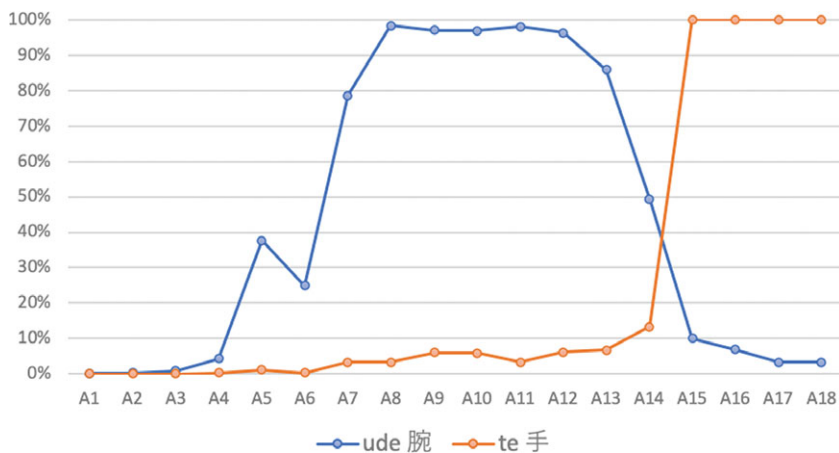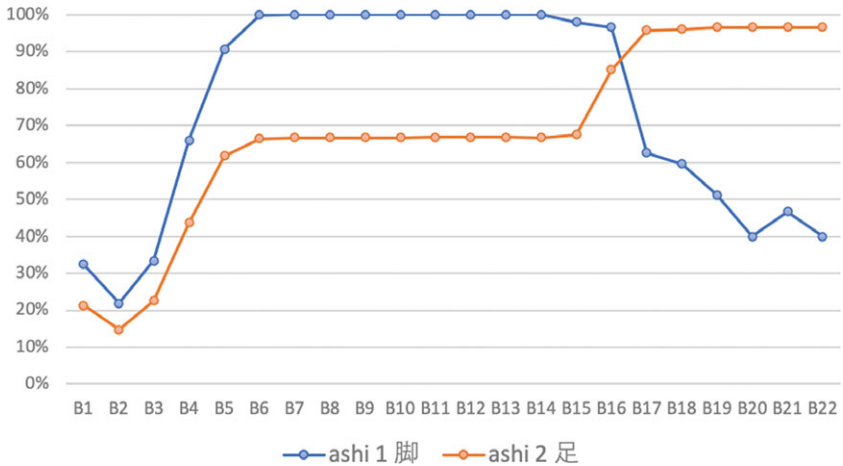

Chart 4

591

Chart 5

conclude that *ude* 腕 and *te* 手are not hierarchically dependent on one another and are at the same level of partonomic depth (i.e., one is not part of the other).

A one-way ANOVA revealed no significant effect of order of stimulus presentation (random, top-down, bottom-up) on the pixel density of A15 ($F(2,27) = 1.58$, $p > .225$). The β-diversity index for the coloring performances was 1.83 for *ude* 腕 and 1.11 for *te* 手. Thus, the semantic extensions of *ude* 腕 and *te* 手 are quite robust and not subject to interpersonal variation in Japanese.

As mentioned in Section 2.2, there are two kanji for *ashi* in Japanese (脚 and 足). We tested both characters, given that there was insufficient evidence to exclude either one, on the basis that, for example, one was more basic than the other. Chart 5 illustrates the semantic extensions of these two lower limb terms.

The distribution of pixel densities was quite different from the 'density overlap' pattern found between the semantic extensions of *bras* and *avant-bras* in French (Chart 2), and which can be considered typical of a hierarchical relation (e.g., part–whole). The pattern illustrated in Chart 5 was also quite different from the clear-cut 'density reversal' pattern observed in the comparison of *ude* 腕 - *te* 手 (Chart 4), *jambe–pied* (Chart 3), or *bras–main* (Chart 1).

In particular, there was overlap in pixel density for the two terms in B6 to B15, suggesting that Japanese speakers did extend the meaning of *ashi 2* 足 to this area of the lower limb but on average less so than the semantic extension of *ashi 1* 脚. There was also a density reversal between B15 and B17, with lower amplitude than the cases reported above (e.g., *jambe–pied*) but still quite noticeable. In other words, only *some* Japanese speakers considered that the meaning of *ashi 2* 足 extended to zones B6–B15 (i.e., the English *leg*,

592

intuitively), whereas *all* considered that the meaning of *ashi 1* 脚 did. More-over, only *some* Japanese speakers considered that the meaning of *ashi 1* 脚 did not extend to zones B17–B22 (i.e., the English *foot*, intuitively), whereas *all* considered that *ashi 2* 足 did. The question is, then, whether the pixel density reversal was significant or not. These two terms are homophones and both refer to the lower limb, therefore we might predict that they have relatively similar meanings and that the differences observed in the descriptive statistics are not significant. To assess whether this null hypothesis could be rejected, we conducted paired samples *t*-tests on the measures of pixel densities of the two zones that are adjacent to the density reversal (B15 and B17). These tests revealed significant differences for both B15 ($t = -4.04$, df = 34.77, *p*-value < .001) and B17 ($t = 3.59$, df = 29.45, *p*-value = .001), allowing us to reject the null-hypothesis and conclude that *ashi 1* 脚 *ashi 2* 足 have different semantic extensions.

We compared whether the area that the Japanese speakers colored in for *ashi 1* 脚 was affected by which stimulus booklet they received with a one-way ANOVA with order of presentation (random, top-down, bottom-up) as the independent variable and the pixel density of the colored anatomical zone B17 as a dependent variable. This analysis revealed no significant effect of presen-tation order on the pixel density of B17 ($F(2,27) = 1.58, p > .225$). Another one-way ANOVA was performed for *ashi 2* 足 and likewise revealed no significant effect of presentation order on the pixel density of B17 ($F(2,27) = 1.19, p > .829$). The β-diversity index for the coloring performances was 1.54 for *ashi 1* 脚 and 1.17 for *ashi 2* 足.

In terms of written language resources, Japanese has two distinct words for the upper limb that do not stand in a hierarchical relation with each other, and that semantically mark the wrist joint boundary. In contradiction to Majid and van Staden's (2015, p. 573) conclusion that Japanese does not have a word for 'foot', the present findings suggest that Japanese also has two distinct words for the lower limb that mark the ankle joint boundary. Whether Japanese speakers use both characters is a question that cannot be addressed by the body coloring task method and remains for future research.

3.1.3. *Indonesian*

Chart 6 illustrates the semantic extensions of two Indonesian upper limb terms *lengan* and *tangan*. The comparison of pixel density distribution across ana-tomical zones for *lengan* and *tangan* revealed a pattern comparable to that of the French *bras* and *avant-bras* (compare Chart 6 and Chart 2). The semantic extensions of both terms overlapped on some anatomical zones (A8–A12), before the pixel density of *lengan* declined in A13 and A14, and dropped off at
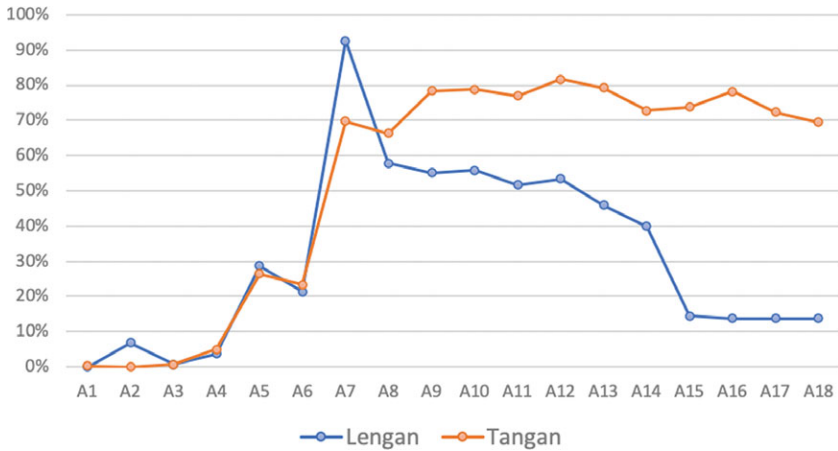
593

Chart 6

A15. This overlapping pattern of distribution suggests that *lengan* is hierarchically related to *tangan* (e.g., *lengan* is part of *tangan*).

Separate one-way ANOVAs revealed no significant effect of presentation order (random, top-down, bottom-up) on the pixel density of A15 for *lengan* ($F(2,27) = 1.58$, $p > .224$) and for *tangan* ($F(2,27) = 0.30$, $p > .742$). The β-diversity indices for the coloring performances of both terms were relatively higher (*lengan*: 3.23; *tangan*: 2.40) than for the French and Japanese performances, suggesting that the meanings are more subject to intrapersonal variation. Although we were careful with the recruitment of Indonesian participants, as detailed in Section 2.1, the multilinguistic environment of Indonesia introduces a multitude of other factors that could potentially impact the robustness of the speaker's semantic knowledge (e.g., language contact with other local languages, speaker's parents' L1, etc.). This ecological specificity could explain why the diversity is relatively higher than for French and Japanese, and should be considered in future research.

Based on the elicitation task that we used to select body part terms for inclusion in the body coloring task, there are no words other than *kaki* in Indonesian that refer to part of the lower limb without being hierarchically related to *kaki* (e.g., *pergelangan kaki* 'ankle'). Chart 7 illustrates the semantic extension of *kaki* in Indonesian.

A one-way ANOVA revealed no significant effect of presentation order (random, top-down, bottom-up) on the pixel density of B20 ($F(2,27) = 2.55$, $p > .096$). We selected B20 as the dependent variable because it was the zone where the average proportion of pixel density was more likely to change, as it did in French (Chart 3). The β-diversity index for the coloring performances of

594
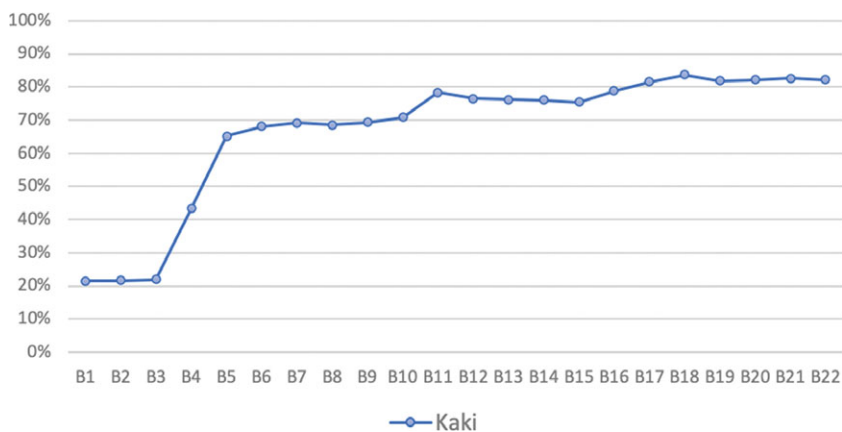
Chart 7

kaki was 1.26. We can therefore conclude that the semantic extension of *kaki* is quite robust in Indonesian.

In sum, *lengan* can now safely be ruled out as a term whose semantic extension is hierarchically unrelated to *tangan*, and we can thus conclude that Indonesian has a term (*tangan*) whose meaning extends to both sides of the wrist joint boundary. Similarly, Indonesian also has a term (*kaki*) whose meaning extends to both sides of the ankle joint boundary.

3.2. THE SEMANTIC EXTENSIONS OF BODY PART TERMS ACROSS LANGUAGE GROUPS

In this section, the semantic extensions of the selected body part terms are compared across languages. To address RQ2, we look for diversity as well as cross-linguistically shared patterns. All supporting ANOVA and post-hoc Bonferroni tests are reported in Tables 2–5 of 'Appendix 2'.

3.2.1 *Upper limb*

As can be seen in Chart 8, the distributions of pixel density for *bras* (French) and *ude* 腕 (Japanese) were very similar from A1 to A18. The distribution of pixel density for *tangan* (Indonesian) was comparable to that of *bras* and *ude* 腕 from A1 to A13. However, the meaning of *tangan* also extended to zones A14–A18, whereas the pixel density for *ude* and *bras* decreased at the 'wrist joint' zones A13/A14 (i.e., the *palmar* and *dorsal carpal ligament*, respectively).

595
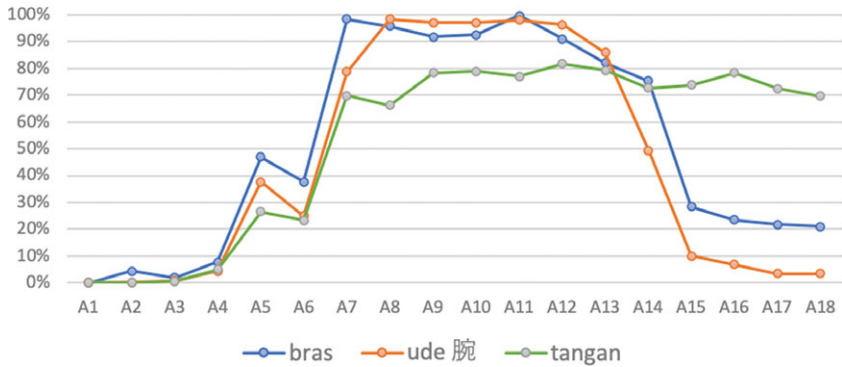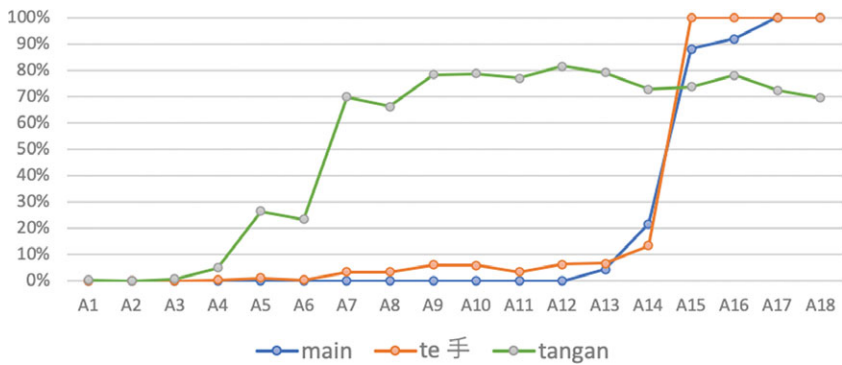
Chart 8

Chart 9

To test whether the differences in pixel density reported for A1–A18 were statistically significant, we conducted separate one-way ANOVAs for each of the 18 anatomical zones, with language (French, Japanese, Indonesian) as the independent variable (Table 2). These analyses revealed a significant effect of language on pixel density for A5, A7–A11, and A15–A18. Post-hoc Bonferroni tests further revealed that Indonesian *tangan* differed from French *bras* (A5, A7), Japanese *ude* 腕 (A9, A10), and from both (A8, A11, A15–A18). These analyses also confirmed that French *bras* and Japanese *ude* 腕 did not differ.

The distributions of pixel density for *main* (French) and *te* 手 (Japanese) appeared to be very similar to each other and quite different from that of *tangan* (Indonesian), as illustrated in Chart 9.

To test whether these observed differences in pixel density were statistically significant, we once again conducted a separate one-way ANOVA for each of the 18 anatomical zones, with language (French, Japanese, Indonesian) as the

596

independent variable (Table 3). These analyses revealed a significant effect of language on pixel density for A3–A18. Post-hoc Bonferroni tests further revealed that Indonesian *tangan* differed systematically from Japanese *te* 手 and from French *main* (with the exception of A15 and A16). These analyses also confirmed that French *bras* and Japanese *ude* 腕 did not differ.

To summarize, the following generalizations can be inferred from the results of the upper limb coloring task in French, Indonesian, and Japanese:

1. French and Japanese use the same type of naming system to refer to the upper limb, which is characterized by:
   a. The absence of a general term for the whole upper limb;
   b. The presence of two distinct terms that are semantically contiguous and hierarchically independent of each other;
   c. Respecting 'the wrist joint boundary rule': the semantic extensions of the two distinct terms do not extend beyond an area of the human body that is constituted by the P A L M A R and D O R S A L C A R P A L L I G A M E N T S (i.e., 'the wrist joint').
2. Indonesian uses a distinct naming system for the upper limb, which is characterized by:
   a. The presence of a general term for the whole upper limb;
   b. The absence of two distinct terms that are semantically contiguous and hierarchically independent of each other;
   c. Violating 'the wrist joint boundary' rule: the semantic extension of the upper limb term extends beyond the *palmar* and *dorsal carpal ligaments*.
3. All three languages respect 'the shoulder joint boundary rule': French, Indonesian, and Japanese do not have a general upper limb term that extends to the area of the upper limb that is constituted by the *clavicula*, *platysma*, and *deltoid* (i.e., the shoulder joint).

In sum, in our limited sample of languages, we found diversity in the naming system of the upper limb body parts (1. vs. 2.) and a cross-linguistically shared pattern (3).

3.2.2. *Lower limb*

As can be seen in Chart 10, the distribution of pixel density for *jambe* (French) and *ashi 1* 脚 (Japanese) was relatively similar, albeit with some notable differences (e.g., B17–B22). The distribution of pixel density for *kaki* (Indonesian) was comparable to that of *jambe* and *ashi 1* 脚 (e.g., B1–B4) but also quite distinct (e.g., B18–B22).

To test whether these observed differences in pixel density were statistically significant, we conducted a separate one-way ANOVA for each of the
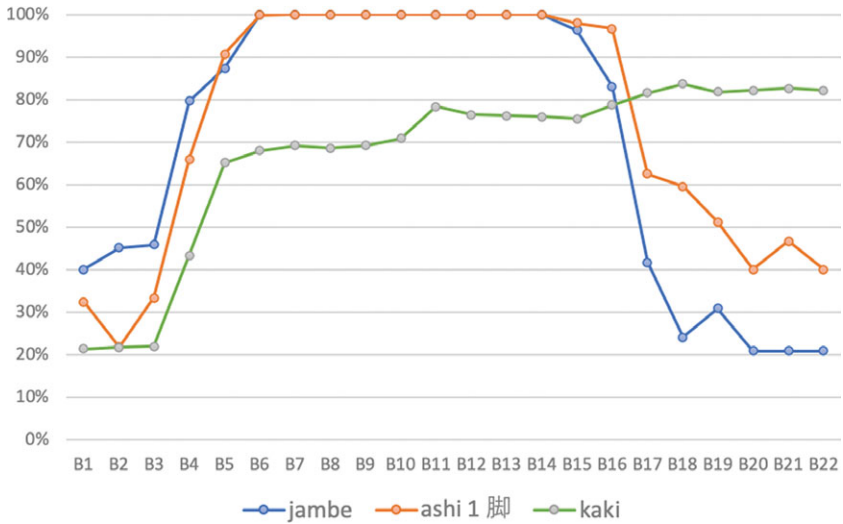
Chart 10

22 anatomical zones, with language (French, Japanese, Indonesian) as the independent variable (Table 4). These analyses revealed a significant effect of language on pixel density for every zone but B16 (i.e., *superior extensor retinaculum*, a group of tendons that extends from the ankle to the heel bone). Post-hoc Bonferroni tests further revealed that the pixel density of every lower limb anatomical zone for Indonesian *kaki* differed from that for French *jambe*, with the exception of B16. Similarly, pixel density differed between Japanese *ashi1* 脚 and Indonesian *kaki*, with the exception of B2 and B16–B18. Finally, post-hoc Bonferroni tests revealed that there was, overall, no difference in pixel density between Japanese *ashi1* 脚 and French *jambe*, with the exception of B2–B4 and B18.

As can be seen in Chart 11, *pied* (French), *ashi 2* 足 (Japanese), and *kaki* (Indonesian) displayed a relative overlap in pixel density from B19 to B22. French *pied*, however, appeared to have a distinct semantic extension from that of *ashi 2* 足 and *kaki* in pixel density for B1–B18. In contrast, the pixel density of B1–B16 largely overlapped for *ashi 2* 足 and *kaki*.

We conducted a separate one-way ANOVA for each of the 22 anatomical zones to test whether these differences are significant with language (French, Japanese, Indonesian) as the independent factor (Table 5). A significant effect of language on pixel density for every zone. Post-hoc Bonferroni tests further revealed the pixel density of French *pied* and Indonesian *kaki* to differ for almost every zone. The pixel density of French *pied* and Japanese *ashi 2* 足

598
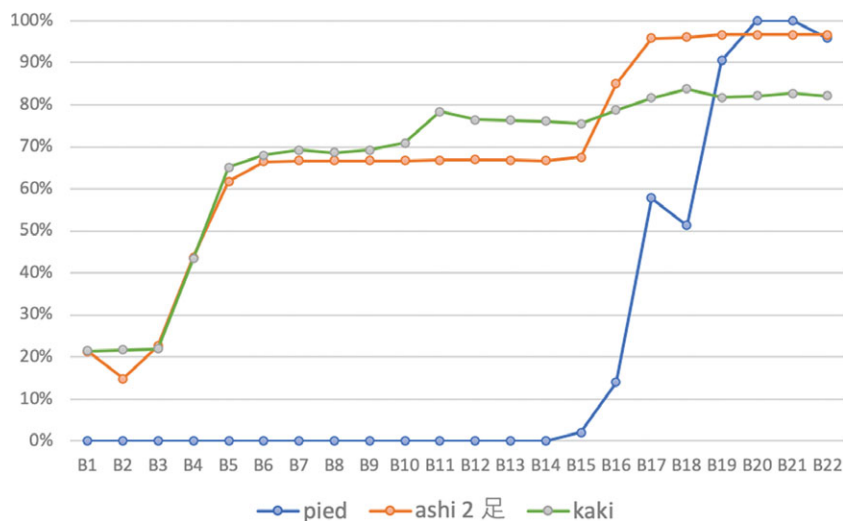
Chart 11

differed for B1–B18, but not for B19–B22, such that there was some semantic overlap between the two terms. The pixel density for Japanese *ashi 2* 足 and Indonesian *kaki* did not differ for most anatomical zones (B1–B18), suggesting substantial semantic overlap between the two terms.

On the one hand, Japanese and French can be grouped together on the basis of having a term that extends to a segment of the lower limb that stops at the ankle joint (*jambe* and *ashi 1* 脚, respectively), in contrast to Indonesian, which has a term which violates the 'ankle joint boundary rule' (*kaki*). On the other hand, Japanese also has a term (*ashi 2* 足) that semantically overlaps with both Indonesian *kaki* (from B1 to B16) and with French *pied* (from B19 to B22). In terms of linguistic systems and written language resources (in contrast to language use) it is, at this point, difficult to force Japanese into either the French type of naming system, which has two distinct, contiguous, and hierarchically independent terms for the lower limb, or the Indonesian type, which has one term that extends to the whole lower limb. We will thus refrain from making generalizations on the Japanese lower limb naming system but can, however, infer from our overall results the following generalizations:

1. French uses a type of naming system to refer to the lower limb that is characterized by:
    a. The absence of a general term for the whole lower limb;
    b. The presence of two distinct terms that are semantically contiguous and hierarchically independent of each other;

599

    c.   Respecting 'the ankle joint boundary rule': the semantic extensions of the two distinct terms do not extend beyond an area of the human body that is constituted by the bottom of the T I B I A and F I B U L A, the T A L U S, M A L E O L L U S, and C A L C A N E U S (i.e., 'the ankle joint').

2.   Indonesian uses a distinct naming system for the lower limb, which is characterized by:

    a.   The presence of a general term for the whole lower limb;

    b.   The absence of two distinct terms that are semantically contiguous and hierarchically independent of each other;

    c.   Violating 'the ankle joint boundary' rule: the semantic extension of the lower limb term extends beyond the *tibia* and *fibula*, the *talus*, *malleolus*, and *calcaneus* (i.e., 'the ankle joint').

3.   All languages respect 'the hip joint boundary rule': French, Indonesian, and Japanese do not have a general lower limb term that extends to the upper lower limb area (i.e., the hip joint).

The cross-linguistic analyses reported in this section illustrate the comparative potential and etic validity of the method we employed in this study. The anatomical grid provides a non-linguistic and culturally neutral yardstick against which we can compare the semantic extension of body part terms across languages. This solves the problem of having to use the emic mini-language of NSM and the pitfalls that come with it. Moreover, this method also allows us to sustain the etic perspective in reporting our results, as there is no need to turn to English as a metalanguage to conclude that 'language X has, or does not have a term for Y', where Y is often reported as an English body part term that inevitably comes with its own language-specific, emic perspective.

3.3. LINGUISTIC MEANING VS NON-LINGUISTIC CATEGORIZATION

The results reported in Sections 3.1 and 3.2 indicate that Indonesian uses a type of body part naming systems that violates 'the wrist joint boundary' rule (with the term *tangan*) and the 'ankle joint boundary' rule (with the term *kaki*), according to which all languages of the world should have distinct and hierarchically independent terms to refer to the body segments situated on each side of the joint boundary (e.g., *main* and *bras* in French). However, *pace* (Wierzbicka, 2007, p. 29) defending that "the most reliable evidence for the presence of such a concept is the presence of a word" and *sec*. Majid (2015, p. 376) answering that "not all concepts are reflected in the lexicon", the fact that Indonesian violates the joint boundary rule in language does not entail that this is necessarily the case in thought. To paraphrase what Goldin-Meadow and Alibali (2013, p. 269) write about gesture, the coloring performances of

600

participants "provides a second window onto the speaker's thoughts, offering insight into those thoughts that cannot be found in speech". More specifically applied to our study, we propose that a discontinuous coloring performance can be interpreted as an index for conceptual discontinuity. For example, in the 90 coloring performances of terms for the upper limb in the three languages, not one participant interrupted their coloring in the middle of the forearm, hence suggesting that this is not a conceptual boundary.

As detailed in Section 2.6, the coloring performances of each participant were coded as discontinuous if done in two clearly distinct sequences. These performances revealed that a substantial number of Indonesian participants stopped at the wrist joint (26 out of 30) and at the ankle joint (21 out of 30) when coloring the regions of the body that corresponded to the meaning of *tangan* and *kaki*, respectively. The coded drawings from the French and Japanese speakers likewise showed that a majority of participants colored upper and lower limbs in two discontinuous segments.

In sum, the following generalization can be inferred from the results of this analysis: the Indonesian body part naming system does not align with the Indonesian body categorization in thought. More broadly, the way we speak of the body does not necessarily determine the way we think about it.

4. General discussion

The study reported in the present paper address a number of methodological and theoretical questions that are central to semantic typology. Methodologically, we controlled for factors that could potentially skew the results of the body coloring task: the multilinguistic environment of participants, their expertise in human anatomy, and the order of presentation of the stimuli. We also implemented an etic perspective at every step of the procedure: in the design of the stimuli, in the coding and analysis of the data, and in the interpretation of results.

Wierzbicka argues that there are no lexical gaps for the concept of 'hand' due to its universally shared experiential salience, whereas Majid and van Staden (2015, pp. 573–574) claim that: "in Indonesian there is no everyday word that corresponds to hand." However, reporting this result as a lexical gap can be problematic. Concluding that 'language X does not have a word for Y' inherently introduces the emic bias of Y. In the case where Y is the English body part term *hand* and X is the Indonesian language, then yes, from an Anglocentric perspective, Indonesian does not have a word for *hand*. But reversing the emic perspective – where Y is the Indonesian body part term *tangan* and X is the English language – then it is English that does not have a word for the whole upper limb. Using English as a metalanguage is inevitable in academic communication, but a guiding principle in semantic typology should be to limit the

601

introduction of culture-specific biases of the English language to the absolute minimum. We have taken a step further towards that goal by reporting the results of our study as follows: *Indonesian belongs to a type of body lexicons that is characterized (among other criteria listed above) by the presence of a general term for the whole upper limb, and the absence of two distinct upper limb terms that are semantically contiguous and hierarchically independent of each other.* The progress is humble, but not trivial if the etic perspective is to prevail in methods of semantic typology, as we argue it should.

Theoretically, the present findings have a bearing on the interaction of language and thought. As pointed out in the 'Introduction', conflating language with thought is a major weakness of Wierzbicka's (2007) NSM approach. Although there is a large body of evidence for a close interaction between language and thought, this interaction does not equal conflation, as a strong version of the Whorfian hypothesis would imply:

> We dissect nature along lines laid down by our native languages. […] [T]he world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds – and this means largely by the linguistic system in our minds. (Whorf, 2013, p. 213)

In the domain of metaphors, some metaphorical expressions appear to be part of a system that aligns with a system of thoughts (e.g., Casasanto & Dijkstra, 2010) and that can influence reasoning (e.g., Thibodeau & Boroditsky, 2011). But there are also sets of linguistic metaphors that do not align with mental ones (e.g., Casasanto & Bottini, 2014) and that even contradict them (e.g., de La Fuente, Casasanto, Román, & Santiago, 2015). One of the main contributions of the present paper has been to test the 'language reflects thinking' hypothesis in the domain of body categorization. We found that there is linguistic diversity in naming body parts across three unrelated languages, but that this diversity does not determine variation in thought. Recent findings (Knight, Bremner, & Cowie, 2020) likewise demonstrate that body lexicons may not shape perception, as having a single term for the entire upper limb (in Croatian) does not impact the boundary effect of the wrist joint in tactile perception.

The study reported in the present paper also allowed a cross-linguistically shared pattern to emerge. Namely, none of the upper limb terms in French, Indonesian, or Japanese extended to the shoulder joint, and none of the lower limb terms extended to the hip joint. This finding can be interpreted as evidence for the role of visual discontinuities and functional salience of certain body parts across body lexicons. Taken together, our data support Majid and van Staden (2015, p. 587): "[s]peakers are clearly attuned to visual

602

discontinuities, but the specific discontinuities they pay attention to are language-specific."

In order to better understand the interaction of language, thought, and perception and to make testable hypotheses, we need a model and a coherent theoretical framework that has greater explanatory potential than a series of isolated findings. This is indeed dearly needed as its absence leaves room for criticism like Wierzbicka (2007, p. 17), who denounces the "theoretical vacuum" and ill-suited methodology of the body coloring task to articulate "a given word's range of use". Three comments can be made about Wierzbicka's critique. First, the body coloring task has never claimed to elicit "the range of use" of body part terms; it maps the meaning of semantic systems, or body lexicons (i.e., language SYSTEMS not language USE, LANGUE not PAROLE). Second, the NSM mini-language is not drawn from authentic corpus-based examples and is therefore, arguably, as decontextualized and inadequate to map the "word's [situated] range of use" as the body coloring task. Third, and more broadly, we could make a case that the Wierzbicka (2007) vs. Enfield et al. (2006) debate goes beyond disagreement on diversity vs. universals and is rooted in a deeper disagreement on the ontology of language.

Blomberg & Zlatev (in press) have indeed argued that the strongly divergent interpretations of linguistic relativity "emanate from radically different under-standings of ontologically loaded concepts like *language* and *thought*". Ulti-mately, the apple of discord is tossed at the conflation of three different levels of language: the level of situated use, the level of sedimented conventions, and the level of prelinguistic motivations. We briefly describe these three levels of meaning-making as integrated in THE MOTIVATION AND SEDIMENTA-TION MODEL (Devylder & Zlatev, 2020; Stampoulidis, Bolognesi, & Zlatev, 2019) before showing how the model brings the present debate under a new light.

Making a doctor's appointment over the phone is an example of a common social interaction where a speaker uses body part terms to serve a specific communicative purpose. Take the case of a patient who has some kind of pain located in the region of the metacarpal bones. Using the 'right word' to describe what or where it hurts is crucial for the accuracy of the initial diagnosis. If that patient is a L1 adult speaker of French, they have acquired its language-specific conventions that limit the meaning of *main* 'hand' to a segment of the body that extends from the wrist to the tip of the fingers. The patient will thus most likely describe their pain as *j'ai mal à la main* 'my hand hurts', but not as *j'ai mal au bras* 'my arm hurts' because the semantic extension of *bras* stops at the wrist joint, as evidenced in this paper. This level of the meaning-making process, where a speaker needs to spontaneously adapt their message to fit a specific situation, is both different from the level of semantic systems, and different from the level of non-linguistic experiences. We could

603

call this level the SITUATED LEVEL of meaning-making, where expressions are subject to interpretation and are highly dependent on the immediacy of the situation. But meaning cannot just be spontaneous language use: the immediacy of the situation that leads the patient to describe their hand injury sits upon deeper levels of the meaning-making process. The doctor can only understand the patient's spontaneous linguistic expression (e.g., *my hand hurts*) if they both share a common set of linguistic conventions (e.g., speak the same language). The *j'ai mal à la main* 'my hand hurts' expression is motivated by the hundreds of years of historical stabilization of the body part lexicon in the French language that sedimented the distinction between *main* and *bras*. This level of relatively stable linguistic conventions, which are not dependent to the communicative context, and where the semantic extensions of body part terms are set within a group of speakers, can be called the SEDIMENTED LEVEL. In addition to the results of our study, the literature on the semantic categorization of the body shows that the semantic segmentation of the body into parts is free to vary across languages WITHIN THE CONSTRAINTS IMPOSED BY THE EMBODIED LEVEL. In other words, languages of the world vary in the marking of perceptual boundaries (e.g., the French type vs. the Indonesian type) but they do not mark a semantic boundary where there is no perceptual boundary (e.g., the middle of the forearm). This suggests that culture-specific conventions do not arbitrarily pop up *ex nihilo* at some point in time in Western Europe, South-East Asia, and Far-Eastern Asia. They do share some common patterns, which implies a common denominator for French, Indonesian, and Japanese speakers. We claim that the common denominator is the experience of the human body, conceived broadly (e.g., the tactile perceptual experience of the wrist joint boundary; the functional salience of hands, etc.). This set of language-specific conventions is thus motivated by a non-linguistic and deeper level of the meaning-making process that can be coined the EMBODIED LEVEL of meaning.

We therefore propose that the meaning-making process is dynamic and consists of at least three different and tightly interacting levels: the embodied level, the sedimented level, and the situated level. This dynamic process is put in motion by two main operations. The MOTIVATION operation links primarily the embodied and situated levels in an 'upward' direction: experiential and cognitive processes are necessary to construct context-specific meaning at the situated level (e.g., in our example, the perceptual and conceptual distinction between hands and arms for the meaning of *mains* and *bras* in the specific context of a doctor's appointment). In a second step, the SEDIMENTATION operation transforms situated sign activities 'downward' into relatively stable structures through frequency of use. These sedimented structures (e.g., the semantic extension of body part terms in a given language) are thus never fully novel and creative, as they also presuppose sedimented norms
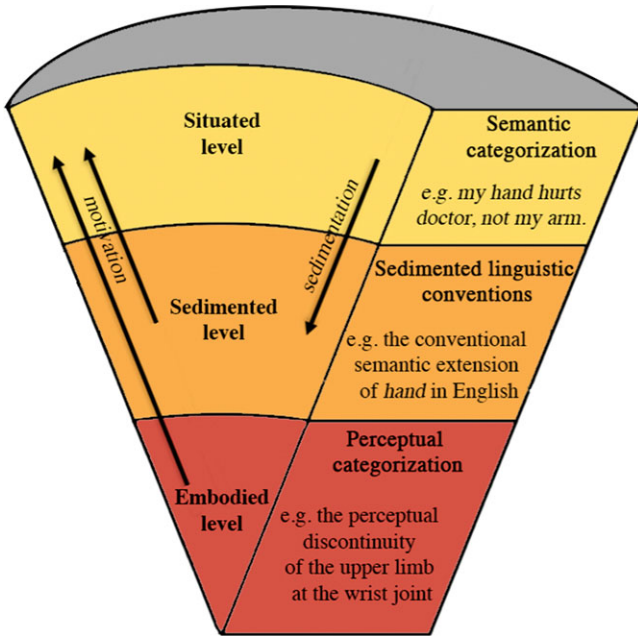
Fig. 4: The motivation and sedimentation of the semantic category of the body.

(Blomberg & Zlatev, in press). Figure 4 is an illustration of a simplified version of the Motivation and Sedimentation Model applied to the meaning-making process that has sedimented the semantic extensions of body part terms that we mapped with the body coloring task.

According to this model, the fact that the meaning of body part terms both vary (as evidenced by our study and by the Special Issue of *Language Sciences*) and share common patterns across languages (as argued by Andersen, 1978; Brown 1976; or Wierzbicka, 2007) is not paradoxical: it is the result of two different steps of the meaning-making process that occur at distinct but tightly interwoven levels. This model also illuminates the phenomenon that we observed, where a majority of Indonesian speakers marked a pause at the wrist and ankle joints in their coloring performances of *tangan* and *kaki*, respectively. Here, we arguably caught on camera a step of the meaning-making process where these Indonesian speakers entered in a dialogical negotiation with the two conflicting body categorizations: one that sets a perceptual boundary at the embodied level (i.e., the boundary effect of the joints of tactile perception), and one that collapses the wrist joint boundary at the sedimented level (i.e., the Indonesian body part lexicon).

The Motivation and Sedimentation Model puts the Wierzbicka (2007) vs. Enfield et al. (2006) debate under a new light: the two sides seem to disagree in part because they are making generalizations about different levels of language. First, neither side is concerned with the situated level, because neither the NSM nor the body coloring task method allows investigating language in use. Second, Wierzbicka conflates the sedimented level (i.e., body part lexicons) and the embodied level (i.e., non-linguistic experience) when she grounds her argument for language universals in the universally shared concept of 'hand' in experiential salience. Third, Enfield et al.'s claim for linguistic diversity of body part lexicons does not inform the meaning-making process happening at the embodied level of language (as critiqued by Wierzbicka) but at the sedimented level of language-specific conventions.

5. Conclusion

The contribution of this paper to the study of meaning aimed to address methodological and theoretical issues with the help of empirical findings. We proposed a development of methods that allow the study of the meaning-making process at work in the categorization of the human body at a fine level of granularity and with considerable cross-linguistic validity. We showed the potential of these methods to provide solid empirical evidence. The broad contribution of this paper consisted of supporting the existence of linguistic diversity in the categorization of the body across languages with an adequate methodology and promising theoretical framework. Finally, we contributed to disentangling the linguistic diversity of body part lexicons vs. universals debate and more broadly conclude that linguistic diversity is not linguistic relativism, in the sense that different categorizations of the body in language do not necessarily determine different categorizations of the body in thought.

REFERENCES

Andersen, E. S. (1978). Lexical universals of body-part terminology. In J. H. Greenberg, C. A. Ferguson & E. A. Moravcsik (eds), *Universals of human language: word structure* (pp. 335–368) (Language Science and National Development Series). Stanford, CA: Stanford University Press.

Blomberg, J. & Zlatev, J. (in press). Metalinguistic relativity: Does one's ontology determine one's view on linguistic relativity? *Language and Communication*.

Brown, C. H. (1976). General principles of human anatomical partonomy and speculations on the growth of partonomic nomenclature 1. *American Ethnologist* **3**(3), 400–424.

Burenhult, N. (2006). Body part terms in Jahai. *Language Sciences* **28**(2/3), 162–180.

Casasanto, D. & Bottini, R. (2014). Mirror reading can reverse the flow of time. *Journal of Experimental Psychology*. *General* **143**(2), 473–479.

Casasanto, D. & Dijkstra, K. (2010). Motor action and emotional memory. *Cognition* **115**(1), 179–185.

606

de La Fuente, J., Casasanto, D., Román, A. & Santiago, J. (2015). Can culture influence body-specific associations between space and valence? *Cognitive Science* **39**(4), 821–832.

de Vignemont, F., Majid, A., Jola, C. & Haggard, P. (2009). Segmenting the body into parts: evidence from biases in tactile perception. *Quarterly Journal of Experimental Psychology* **62**(3), 500–512.

Devylder, S. & Zlatev, J. (2020). Cutting and breaking metaphors of the self and the Motivation & Sedimentation Model. In Annalisa Baicchi (ed.), *Figurative meaning construction in thought and language* (pp. 253–282). Amsterdam/Philadelphia, PA: John Benjamins.

Enfield, N. J., Majid, A. & van Staden, M. (2006). Cross-linguistic categorisation of the body: introduction. *Language Sciences* **28**(2/3), 137–147.

Goldin-Meadow, S. & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology* **64**, 257–283.

Knight, F. L. C., Bremner, A. J. & Cowie, D. (2020). Does the language we use to segment the body, shape the way we perceive it? A study of tactile perceptual distortions. *Cognition* **197**, e104127.

Knight, F. L. C., Cowie, D. & Bremner, A. J. (2017). Part-based representations of the body in early childhood: evidence from perceived distortions of tactile space across limb boundaries. *Developmental Science* **20**(6), e12439.

Knight, F. L. C., Longo, M. R. & Bremner, A. J. (2014). Categorical perception of tactile distance. *Cognition* **131**(2), 254–262.

Lausberg, H. & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES−ELAN system. *Behavior Research Methods* **41**(3), 841–849.

Levinson, S. C. (2006). Parts of the body in Yélî Dnye, the Papuan language of Rossel Island. *Language Sciences* **28**(2/3), 221–240.

Majid, A. (2006). Body part categorisation in Punjabi. *Language Sciences* **28**(2/3), 241–261.

Majid, A. (2010). Words for parts of the body. In B. C. Malt & P. M. Wolff (eds), *Words and the mind: how words capture human experience* (pp. 58–71). New York/Oxford: Oxford University Press.

Majid, A. (2015). Comparing lexicons cross-linguistically. In J. R. Taylor (ed.), *The Oxford handbook of the Word* (pp. 364–379). Oxford: Oxford University Press.

Majid, A. & van Staden, M. (2015). Can nomenclature for the body be explained by embodiment theories? *Topics in Cognitive Science* **7**(4), 570–594.

Plato (1925). *Plato in twelve volumes: Vol. 9*, H. N. Fowler, trans. Cambridge, MA: Harvard University Press; London: William Heinemann.

Shen, G., Smyk, N. J., Meltzoff, A. N. & Marshall, P. J. (2018). Neuropsychology of human body parts: exploring categorical boundaries of tactile perception using somatosensory mismatch responses. *Journal of Cognitive Neuroscience* **30**(12), 1858–1869.

Sneddon, J. N. (2006). *Colloquial Jakartan Indonesian* (Pacific Linguistics 581). Canberra: Pacific Linguistics.

Sneddon, J. N., Adelaar, K. A., Djenar, D. N. & Ewing, M. (2012). *Indonesian: a comprehensive grammar*. London: Routledge.

Stampoulidis, G., Bolognesi, M. & Zlatev, J. (2019). A cognitive semiotic exploration of metaphors in Greek street art. *Cognitive Semiotics* **12**(1), e20192008.

Stevens, A. M. & Schmidgall-Tellings, A. (2004). *A comprehensive Indonesian–English dictionary*. Athens, OH: Ohio University Press.

Tadmor, U. (2017). Malay–Indonesian. In B. Comrie (ed.), *The world's major languages* (pp. 809–837). London: Routledge.

Terrill, A. (2006). Body part terms in Lavukaleve, a Papuan language of the Solomon Islands. *Language Sciences* **28**(2/3), 304–322.

Thibodeau, P. H. & Boroditsky, L. (2011). Metaphors we think with: the role of metaphor in reasoning. *PloS One* **6**(2), e16782.

van Staden, M. & Majid, A. (2006). Body colouring task. *Language Sciences*, **28**(2/3), 158–161.

Wegener, C. (2006). Savosavo body part terminology. *Language Sciences* **28**(2/3), 344–359.

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon* **21**(2/3), 213–251.

607

Whorf, B. L. (ed.). (2013). *Language, thought, and reality: Selected Writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.

Wierzbicka, A. (2007). Bodies and their parts: an NSM approach to semantic typology. *Language Sciences* **29**(1), 14–65.

Wierzbicka, A. & Goddard, C. (2017). Talking about our Bodies and their Parts in Warlpiri. *Australian Journal of Linguistics* **38**(1), 31–62.

Appendix 1

Description of anatomical zones

Zone	Anatomy
A1	*clavicula + platysma*
A2	*trapezius muscle*
A3	*pectoral major muscle*
A4	*infraspinatus fascia*
A5	*deltoid muscle*
A6	*deltoid muscle*
A7	*biceps brachii*
A8	*triceps brachii*
A9	*brachioradialis + extensor carpi radialis longus and brevis*
A10	*pronator teres + flexor carpi radialis + palmaris longus + flexor digitorum superficialis + flexor carpi ulnaris + flexor digitorum profundus + flexor pollicis longus + pronator quadratus + radius + bicipital aponeurosis*
A11	*olecranon of ulna*
A12	*flexor carpi ulnaris + extensor carpi ulnaris + extensor digitorum + extensor digiti minimi*
A13	*palmar carpal ligament*
A14	*dorsal carpal ligament*
A15	*palm*
A16	*dorsum manus*
A17	*natatory ligament*
A18	*thumb*
A19	*index finger*
A20	*middle finger*
A21	*ring finger*
A22	*little finger*
A23	*thumb*
A24	*index finger*
A25	*middle finger*
A26	*ring finger*
A27	*little finger*
B1	*iliotibial tract + vastus lateralis*
B2	*gluteus maximus*
B3	*sartorius + iliacus + pectineus + adductor longus + gracilis*
B4	*rectus femoris*
B5	*semitendinosus + biceps femoris + semimembranosus*
B6	*suprapatellar bursa*
B7	*vastus medialis muscle*
B8	*subcutaneous prepatellar bursa*
B9	*fibrous joint capsule + iliotibial tract*
B10	*fossa poplitea*
B11	*peroneus longus muscle*

608

Continued

B12	*tibialis anterior muscle*
B13	*soleus muscle*
B14	*gastrocnemius muscle*
B15	*calcaneal tendon*
B16	*ankle*
B17	*inferior extensor retinaculum*
B18	*malleolus*
B19	*heel*
B20	*dorsum pedis*
B21	*planta*
B22	*fifth toe*
B23	*fourth toe*
B24	*third toe*
B25	*second toe*
B26	*big toe*
B27	*fifth toe nail*
B28	*fourth toe nail*
B29	*third toe nail*
B30	*second toe nail*
B31	*big toe nail*
B32	*fifth toe*
B33	*fourth toe*
B34	*third toe*
B35	*second toe*
B36	*big toe*

609

Appendix 2

ANOVA and post-hoc Bonferroni test tables

T A B L E 2. *One-way analysis of variance of pixel density for* bras, ude, tangan *by language group and post-hoc Bonferroni tests*

	One-way ANOVA				Post-hoc Bonferroni test		
Zone	*df-groups*	*df-residuals*	*F*	*p*	*FRE-ƷPN*	*FRE-IND*	*ƷPN-IND*
A1	2	87	0.67	> .0515	–	–	–
A2	2	87	0.93	> .399	–	–	–
A3	2	87	1.71	> .188	–	–	–
A4	2	87	1.51	> .227	–	–	–
A5	2	87	3.54	*.0332	–	*p = .029	–
A6	2	87	1.82	> .167	–	–	–
A7	2	87	6.11	**.0033	–	**p = .0026	–
A8	2	87	16.46	***< .001	–	***p < .001	*** p < .001
A9	2	87	3.90	*.024	–	–	*p = .024
A10	2	87	4.17	*.0187	–	–	*p = .02
A11	2	87	9.02	***< .001	–	***p < .001	***p < .001
A12	2	87	2.64	> .0769	–	–	–
A13	2	87	0.577	> .564	–	–	–
A14	2	87	2.58	> .0816	–	–	–
A15	2	87	29.21	***< .001	–	***p < .001	***p < .001
A16	2	87	37.44	***< .001	–	***p < .001	***p < .001
A17	2	87	34.27	***< .001	–	***p < .001	***p < .001
A18	2	87	30.24	***< .001	–	***p < .001	***p < .001

TABLE 3. *One-way analysis of variance of pixel density for* main, te, tangan *by language group and post-hoc Bonferroni tests*

	One-way ANOVA				Post-hoc Bonferroni test		
Zone	df-groups	df-residuals	F	p	FRE-JPN	FRE-IND	JPN-IND
A3	2	87	4.13	*.0193	–	*p = .044	*p = .044
A4	2	87	9.92	***< .001	–	***p < .001	***p < .001
A5	2	87	21.02	***< .001	–	***p < .001	***p < .001
A6	2	87	16.92	***< .001	–	***p < .001	***p < .001
A7	2	87	55	***< .001	–	***p < .001	***p < .001
A8	2	87	49.92	***< .001	–	***p < .001	***p < .001
A9	2	87	69.66	***< .001	–	***p < .001	***p < .001
A10	2	87	75.49	***< .001	–	***p < .001	***p < .001
A11	2	87	71.94	***< .001	–	***p < .001	***p < .001
A12	2	87	80.41	***< .001	–	***p < .001	***p < .001
A13	2	87	64.11	***< .001	–	***p < .001	***p < .001
A14	2	87	20.19	***< .001	–	***p < .001	***p < .001
A15	2	87	6.79	**.00181	–	–	**p = .0012
A16	2	87	5.25	**.00701	–	–	**p = .0064
A17	2	87	11.55	***< .001	–	***p < .001	***p < .001
A18	2	87	12.38	***< .001	–	***p < .001	***p < .001

TABLE 4. *One-way analysis of variance of pixel density for* jambe, kaki, ashi1 脚 *by language group and post-hoc Bonferroni tests*

Zone	df-groups	df-residuals	F	p	FRE-JPN	FRE-IND	JPN-IND
		One-way ANOVA				Post-hoc Bonferroni test	
B1	2	87	19.69	***< .001	–	***p < .001	***p = .0007
B2	2	87	5.744	**.00454	**p = .0095	*p = .0177	–
B3	2	87	27.2	***< .001	***p = .00094	***p < .001	***p = .00149
B4	2	87	26.86	***< .001	*p = .034	***p < .001	***p < .001
B5	2	87	6.97	**.00156	–	**p = .0064	**p = 0043
B6	2	87	14.23	***< .001	–	***p < .001	***p < .001
B7	2	87	13.93	***< .001	–	***p < .001	***p < .001
B8	2	87	14.41	***< .001	–	***p < .001	***p < .001
B9	2	87	14.18	***< .001	–	***p < .001	***p < .001
B10	2	87	12.98	***< .001	–	***p < .001	***p < .001
B11	2	87	9.52	***.000184	–	***p = .00087	***p = .00087
B12	2	87	10.94	***< .001	–	***p = .00033	***p = .00033
B13	2	87	10.77	***< .001	–	***p = .00037	***p = .00037
B14	2	87	11.17	***< .001	–	***p = .00028	***p = .00028
B15	2	87	8.48	***.000431	–	***p = .0031	***p = .0011
B16	2	87	2.85	.0631	–	–	–
B17	2	87	7.139	**.00135	–	**p = .0009	–
B18	2	87	14.68	***< .001	**p = .0034	***p < .001	–
B19	2	87	11.13	***< .001	–	***p < .001	*p = .032
B20	2	87	14.5	***< .001	–	***p < .001	**p = .002
B21	2	87	14.26	***< .001	–	***p < .001	*p = .011
B22	2	87	14.5	***< .001	–	***p < .001	**p = .002

TABLE 5. *One-way analysis of variance of pixel density for* pied, kaki, ashi 2 足 *by language group and post-hoc Bonferroni tests*

	One-way ANOVA				Post-hoc Bonferroni test		
Zone	df-groups	df-residuals	F	p	FRE-ЈPN	FRE-IND	ЈPN-IND
B1	2	87	27.91	***< .001	***p < .001	***p < .001	–
B2	2	87	8.695	***.000361	*p = .03317	***p = .00025	–
B3	2	87	24.64	***< .001	***p < .001	***p < .001	–
B4	2	87	27.06	***< .001	***p < .001	***p < .001	–
B5	2	87	28.51	***< .001	***p < .001	***p < .001	–
B6	2	87	28.98	***< .001	***p < .001	***p < .001	–
B7	2	87	30.11	***< .001	***p < .001	***p < .001	–
B8	2	87	29.85	***< .001	***p < .001	***p < .001	–
B9	2	87	30.43	***< .001	***p < .001	***p < .001	–
B10	2	87	31.31	***< .001	***p < .001	***p < .001	–
B11	2	87	39.95	***< .001	***p < .001	***p < .001	–
B12	2	87	39.96	***< .001	***p < .001	***p < .001	–
B13	2	87	39.05	***< .001	***p < .001	***p < .001	–
B14	2	87	38.22	***< .001	***p < .001	***p < .001	–
B15	2	87	35.53	***< .001	***p < .001	***p < .001	–
B16	2	87	39.77	***< .001	***p < .001	***p < .001	–
B17	2	87	7.58	***.000924	***p = .00058	–	–
B18	2	87	9.23	***.000231	***p = .00016	*p = .02740	–
B19	2	87	3.3	*.0416	–	–	*p = .038
B20	2	87	5.20	**.00738	–	**p = .010	**p = .042
B21	2	87	5.11	**.00795	–	**p = .011	**p = .045
B22	2	87	3.59	*.0318	–	p = .068	p = .068