What is psycho-collocation?

Preliminary considerations on terminology, definitions and conceptual issues

Daniel Majchrzak

Bachelor of Philosophy, Honours in Language Studies

The Australian National University

December 2025

This thesis is submitted in partial fulfilment of the requirements for the degree of Bachelor or Philosophy with Honours in Language Studies in the College of Arts and Social Sciences.

I hereby declare that, except where it is otherwise acknowledged in the text, this thesis represents my own original work.

All versions of the submitted thesis (regardless of submission type) are identical. The version of the thesis which has been uploaded to the ANU Print Repository additionally integrates comments given by examiners after submission.

Daniel Majchrzak

December 2025 (Updated July 2026)

You talked, but after your talking all the rest remains.

After your talking—poets, philosophers, contrivers of romances—

Everything else, all the rest deduced inside the flesh

Which lives and knows, not just what is permitted.

Czesław Miłosz, *Separate Notebooks* (A. Miłosz, Trans., 1986, p. 25)

Acknowledgement

Firstly, to my supervisor Zhengdao Ye. Thank you for your enduring patience, support, insight and assistance with editing. I looked forward to each of our meetings knowing they would leave me more confident in my next step.

There are a number of other people at ANU who I would also like to thank for their support. Both Rosie Billington and Thomas Nulley-Valdés helped greatly in their role as Honours Convenors and both Jane Simpson and Grace Koch were extremely supportive when it came to managing Honours alongside work as a research assistant. After submission, the version of this thesis that has been uploaded to the ANU Thesis Repository greatly benefitted from comments and criticism from my examiners: Avery Andrews and John Newman.

To my family, thank you for your support which came in various forms including conversations, encouragement, distraction and food. I would also like to acknowledge in particular the encouragement and editing skills of my mum, who was able to see things in what I had written that I was too close to see and who gave her time so freely to support me. Additionally, to Pamballams, who are effectively my family and who bring a joy that I hope to carry into other areas of my life. One ever-important source of this joy was getting breakfast on Thursdays with my mum Annie, sister Becca and later on my (de facto) brother Trell.

To all of the people who, at one time or another, talked to me about what I was writing or some other topic which made me think deeply. Often the scattered thoughts I was having about my thesis would come together after these conversations. In this vein I would particularly like to thank Jack Batten, Joseph Baxter, Connor Brown, Tom Cliff, Becca Majchrzak, Annie O'Rourke and Zhengdao Ye. Similarly, I was helped by email correspondence with Alice Vittrant and Aimée Lahaussois, the latter of whom graciously allowed me to read a manuscript for a submitted article (Lahaussois, In Press).

Lastly, to all the world's beautiful things, but especially to language, which steadily becomes more beautiful the more I learn.

Abstract

Language is central to communicating what we think, sense and feel, with languages differing in the ways they communicate these experiences. This thesis explores *psycho-collocation*, a kind of "polymorphemic expression referring as a whole to a mental process, quality or state" (Matisoff 1986:4). While the term *psycho-collocation* has largely been applied to languages of mainland Southeast Asia, there are a number of apparently similar constructions beyond the region. This thesis extends this research by exploring the possibility of extending *psycho-collocation* to languages in mainland Southeast Asia. Further, it gives a preliminary characterisation of some lexical-semantic aspects of psycho-collocation for a diverse sample of five languages: Dalabon, Ewe, Hakha Lai, Mangap-Mbula and Yucatec Maya.

I first consider the use of the term *psycho-collocation* in mainland Southeast Asia and consider several apparently similar concepts which have been applied outside of this region. With the aim of unifying these various concepts, I then develop a general preliminary definition of psycho-collocation by appealing to the notion of the *experiential construction* in Verhoeven (2007), and Melčuk's (2023) formulation of the *lexemic phraseme*.

With this preliminary definition in hand, I give an overview of the potential for this conception of psycho-collocation in five languages:

Dalabon, Ewe, Hakha Lai, Mangap-Mbula and Yucatec Maya. Using a database of relevant examples from these languages, I then look across the sampled languages to consider whether one could distinguish aspects of “typical” and “atypical” psycho-collocation in these languages. I conclude that a significant portion of the preliminary definition of psycho-collocation can be retained as a kind of comparative concept and that there are several lexemic-semantic aspects of psycho-collocation where “typical” and “atypical” psycho-collocation can be differentiated. I conclude by considering how psycho-collocation fits among other comparative concepts, and suggest considerations and directions for future research.

Table of Contents

Acknowledgement....iii
Abstract....v
Table of Figures....ix
Index of Tables....x
Glossing conventions and abbreviations....xii
1 Introduction....1
2 Literature Review....5
2.1 Why Experiential Constructions?....5
2.1.1 Motivating the study of experiential constructions....5
2.1.2 Overview of experiential constructions....8
2.1.3 Subdomains of the functional domain of experience....14
2.2 Psycho-collocations....18
2.2.1 Broad and narrow psycho-collocation....19
2.2.2 Semantic domains of psycho-collocation....24
2.2.3 Psycho-collocation outside of Mainland Southeast Asia....28
2.3 Phraseology and idioms....30
2.3.1 Psycho-collocations as either idioms or collocations....31
2.3.2 Preliminary definitions of broad and narrow psycho-collocation 34
3 Approach....39
3.1 General characterisation of method and assumptions....39
3.2 Languages considered....42
3.3 Constructing a database of psycho-collocations....44
4 Overviews of Individual Languages....49
4.1 Dalabon....49
4.1.1 Person-parts in Dalabon Psycho-collocations....50
4.1.2 Morphosyntactic aspects of Dalabon psycho-collocations....55
4.2 Ewe....58
4.2.1 Person-parts in Ewe Psycho-collocations....58
4.2.2 Morphosyntactic aspects of Ewe psycho-collocations....62
4.3 Hakha Lai....65
4.3.1 Person-parts in Hakha Lai Psycho-collocations....65

4.3.2 Morphosyntactic aspects of Hakha Lai psycho-collocations.....69
4.4 Mangap-Mbula.....72
4.4.1 Person-parts in Mangap-Mbula Psycho-collocations.....73
4.4.2 Morphosyntactic aspects of Mangap-Mbula psycho-collocations..77
4.5 Yucatec Maya.....79
4.5.1 Person-parts in Yucatec Maya Psycho-collocations.....80
4.5.2 Morphosyntactic aspects of Yucatec Maya psycho-collocations....84
4.6 Overview.....86
5 General analysis toward a refined definition of psycho-collocation.....88
5.1 Psycho-collocations as experiential constructions.....89
5.1.1 Experiential subdomains in the database.....89
5.1.2 Comparison of “core” meanings of experiential subdomains.....95
5.1.3 Beyond experiential subdomains in the database.....103
5.2 Comparison of person-parts in psycho-collocation.....112
5.3 Psycho-collocations as lexemic phrasemes.....124
5.3.1 Bickel’s (1997) typology of the possessive of experience.....125
5.3.2 Constrained selection of lexemic components.....129
5.3.3 Proportion of Bickel type by language and whether psycho-collocations as idioms or collocations.....137
5.3.4 (Non-)compositionality.....139
5.4 A refined definition of psycho-collocation.....142
6 General discussion.....147
6.1 Psycho-collocation as a comparative concept.....147
6.2 Considerations and limitations for future investigations of psycho-collocation.....153
7 Conclusion.....162
Reference List.....164
Appendices.....174
Appendix A. Spearman’s rank correlation between frequency of person-part and number of subdomains in entries with that person-part for each language.....174
Appendix B. Clustering of person-parts by subdomain in each language..175
Appendix C. Contingency table between subdomain and clusters given in Figure 5.4.....177

Table of Figures

Figure 3.1. Components of a basic experiential situation (per Verhoeven 2007, p.52 ..... *12*

Figure 3.2. Subdomains in Becher (2003), Verhoeven (2007), and Malchukov (2015) *15*

Figure 4.1. Number of identified psycho-collocations by language *47*

Figure 4.2. Example database entry *48*

Figure 6.1. Proportion of each subdomain in the database for unique mate-noun pairings *90*

Figure 6.2. Number of co-occurrences between subdomains for each language *92*

Figure 6.3. Number of person parts in unique mate-noun pairings by language. *115*

Figure 6.4. Grouping of person-parts across languages based on experiential subdomains of associated database entries *118*

Figure 6.5. Standardised residuals for chi-squared test of association between clusters given in Figure 5.4 and individual experiential subdomains *120*

Figure 6.6. Summary of the types of possessive of experience construction given in Bickel (1997) *129*

Figure 6.7. Unique noun-mate pairings for each language by types given in Bickel (1997) *138*

Index of Tables

Table 2.1. Members of the basic experiential situation (per Verhoeven 2007, p.52) *12*

Table 2.2. Overview of the functional domain of experience's subdomains following Verhoeven (2007) *16*

Table 2.3. Changes in the use of the term psycho-collocation since Matisoff (1986) *20*

Table 2.4. Differences in the semantic range of psycho-collocation for some publications on MSEA languages *24*

Table 2.5. Terms used outside MSEA which appear closely related to psycho-collocation *28*

Table 3.1. Sources used to gather examples in the database *43*

Table 3.2. Dictionaries used in database *44*

Table 4.1. Dalabon person-parts in database *50*

Table 4.2. Person-parts and subdomains in a sample of Dalabon Psycho-collocations (unique noun-mate pairings) *51*

Table 4.3. Ewe person-parts in database *58*

Table 4.4. Person-parts and subdomains in a sample of Ewe Psycho-collocations (unique noun-mate pairings) *58*

Table 4.5. Hakha Lai person-parts in database *64*

Table 4.6. Person-parts and subdomains in a sample of Hakha Lai Psycho-collocations (unique noun-mate pairings) *65*

Table 4.7. Mangap-Mbula person-parts in database *72*

Table 4.8. Person-parts and subdomains in a sample of Mangap-Mbula Psycho-collocations (unique noun-mate pairings) *74*

Table 4.9. Yucatec Maya person-parts in database *79*

Table 4.10. Person-parts and subdomains in a sample of Yucatec Maya Psycho-collocations (unique noun-mate pairings) *80*

Table 5.1. Proportion of unique mate-noun pairs coded as a single subdomain *92*

Table 5.2. Select predicate meanings from NSM and ValPaL sorted by experiential subdomain *96*

Table 5.3. Potential psycho-collocation identified select experiential meanings *97*

Table 5.4. Examples of person-parts which are used to express experiential meanings without requiring a psycho-mate.....99

Table 5.5. Examples in the database which express interpersonal meanings.....105

Table 5.6. Examples in the database which denote physiological, biochemical or behavioural reactions rather than experiences themselves.....107

Table 5.7. Examples in the database referring to character traits or dispositions 110

Table 5.8. Frequency statistics for clusters in Figure 5.4.....121

Table 5.9. Bickel's (1997) typology of the lexical composition in "possessive of experience" constructions.....126

Table 5.10. Unique noun-mate pairings for each language by types given in Bickel (1997).....136

Glossing conventions and abbreviations

Interlinear glossing follows the following format, where each dot point below represents a line in the interlinear gloss.

- Language name and glottocode where the language is not one of the five that are the focus of this thesis (cf. §3.2) or English
- Untranslated sentence
- Morpheme by morpheme translation
- Free and literal translation of example

The first line contains the language name, and a glottocode where the language is not one of the five that are the focus of this thesis (cf. §3.2). This line is omitted in Chapter 4, where each section focuses on a single language. The second line is the untranslated example sentence indicating morpheme boundaries. The third line is a morpheme by morpheme translation, following the glossing conventions applied in the original source. The final line contains a free and literal translation (if provided), as well as the associated reference. This final line is repeated for some examples where multiple alternate translations are provided.

The abbreviations in each gloss follow those which were used in the original source unless otherwise stated.

1 Introduction

Humans rely a great deal on language for conveying our emotions and inner states. One way of doing so is by involving parts of the body in descriptions, as in the English phrase “my heart sank”. For some languages, researchers have noted that parts of the person are a highly productive means of conveying emotions and inner states. This thesis investigates such cases by extending existing work on *psycho-collocation*, a construction first described in Matisoff (1986) for languages in the Mainland Southeast Asia language area (MSEA) as:

> “[a] polymorphemic expression referring as a whole to a mental process, quality or state, one of whose constituents is a *psycho-noun*, i.e. a noun with explicit psychological reference … The rest of the [psycho-collocation] constrains morphemes (usually action verbs or adjectives) that complete the meaning. This element we call the *psycho-mate*” (Matisoff, 1986 p.4; emphasis in original).

Although psycho-collocations are “especially elaborated” in MSEA, there are analogous constructions which appear across the world’s languages (Enfield, 2021, p.222). This was evident in the original conception of the term, given Matisoff contrasts the strong tendency toward psycho-collocation in MSEA to existent but less phenotypical expressions in several Standard Average European languages (1986:3). While an

analogy has been drawn to psycho-collocation in a number of studies on individual languages outside of MSEA (for example in (Ameka, 1990; Bugenhagen, 1990; Clark, 1989; Fedriani, 2014; Verhoeven, 2007; von Prince, 2017; *inter alia*), psycho-collocation has yet to receive dedicated comparative study beyond MSEA since Matisoff (1986).

This thesis takes steps towards filling this research gap by investigating psycho-collocations across a small but diverse sample of languages: Dalabon, Ewe, Hakha-Lai, Mangap-Mbula, and Yucatec Maya (cf. §3.2). By doing so, it aims to contribute to the understanding of psycho-collocations outside of mainland Southeast Asia and to outline how they might fit into a larger typology of experiential constructions. I do this by investigating psycho-collocations with respect to two research questions in particular, which I outline below.

A first research question relates to the suitability of existing definitions of psycho-collocation for languages both inside and outside of the Mainland Southeast Asia language area (MSEA). The notion of psycho-collocation originates in Matisoff's (1986) study of the phenomenon, which aims to distinguish the construction in languages of the MSEA (Enfield, 2021). Later studies which consider similar phenomena for languages in MSEA refer back to Matisoff's study when defining psycho-collocation (Jaisser, 1990; Mulder, 2021; Oey, 1990; VanBik, 1998; VanBik, 2010; Vittrant, 2013; *inter alia*). However, it is unclear whether studies

which apply the term *psycho-collocation* are discussing the same phenomena as discussed in Matisoff (1986). This problem would likely be compounded for studies mentioning psycho-collocation outside of MSEA (e.g. Fox, 2018; Verhoeven, 2007; von Prince, 2017; Weissmann, 2023; *inter alia*) given their relative dissimilarity. Because of the importance of well-defined comparative concepts for the purpose of investigating linguistic phenomena across languages (Haspelmath, 2010, 2018, 2025), an important objective of this study is to consider the applicability of psycho-collocation for comparison beyond MSEA.

To assist in addressing this first research question, I consider a second research question relating to the comparison of aspects of what can be broadly considered psycho-collocation in the languages under investigation. There is a sizable body of research that notes the different ways that psycho-collocations differ from one another in the languages of Mainland South-east Asia (Bickel, 2004; Clark, 1989; Matisoff, 1986; Van-Bik, 1998; Vittrant, 2013); Vittrant and Watkins 2019; *inter alia*). An important objective of this research is to investigate the extent to which some of these aspects of psycho-collocation as investigated within MSEA can be applied beyond that language area. In particular, I focus on the semantic contribution of the psycho-noun and psycho-mate to the larger meaning of psycho-collocations, on the semantic domain denoted by the psycho-collocation as whole, the structural relationship between the

psycho-noun and the psycho-mate, and the part of speech of the psycho-mate.

To answer these questions this thesis will broadly follow the (functional-) typological approach (cf. Croft, 2002; Fedriani, 2014). Within this broader methodological tradition, this study aligns with a more specific approach taken for semantically oriented typology in particular, as outlined by Evans (Evans, 2004, 2010). Some assumptions taken by these approaches and greater detail on the method taken here are given in Section 3.1

This thesis is structured as follows. Chapter 2 introduces key concepts, discusses differences between studies on psycho-collocation and considers the above research questions in more detail. Chapter 3 introduces the method taken here, the languages sampled and a dataset of psycho-collocations. Chapter 4 discusses psycho-collocation for each of the languages under investigation. Chapter 5 compares the languages discussed in Chapter 4, with the aim of creating a revised definition of psycho-collocation. Chapter 6 places psycho-collocation among other comparative concepts, as well as discussing limitations and future directions. Chapter 7 briefly overviews and concludes the study.

2 Literature Review

The notion of psycho-collocation sits at the intersection of several areas of linguistic research. This chapter aims to outline the conceptual framework employed in this thesis. In Section 2.1 I briefly motivate the topic of study by introducing the body of research concerned with how different languages encode experiencers. I then broaden the scope to the more general notion of experiential constructions. This is followed by Section 2.2, which discusses the origin and applications of the term "psycho-collocation", first as used in Mainland Southeast Asia and then elsewhere. Section 2.3 introduces phraseology as a significant area of research for studying psycho-collocations. I conclude by defining psycho-collocation with respect to the background given in the preceding sections.

2.1 Why Experiential Constructions?

2.1.1 Motivating the study of experiential constructions

There are a large number of studies which are interested in investigating the different ways that languages function when used to convey experiences. Considered in a broad sense, psycho-collocations can be understood as part of this intellectual environment. Here I give examples from two bodies of research as a way of motivating investigation of psycho-collocations. Beginning with English examples

(1-3), note the differences in the ability of the verbs *eat*, *frighten* and *fear* to form verbal compounds as given in (1b), (2b) and (3b) respectively (Grimshaw, 1990 p.15, ex. 18, 20, 28):[1]

(1) a. The tiger eats the man

b. A man-eating tiger

(2) a. God frightens man

b. *A man-frightening god

(3) a. Man fears god

b. A god-fearing man

Grimshaw notes that in English a compound can be formed from action verbs like *eat*, as in example (1). But, as seen in (2), this is not the case for all predicates. Further, the semantic similarity between *fear* in (2) and *frighten* (3), alongside their differences with respect to forming verbal compounds, suggests that there are divisions within the class of English psychological verbs (which includes words like *fear*) that might not be immediately apparent when considering their semantic similarity.

The above examples therefore demonstrate that psychological verbs can be distinguished from so-called 'action verbs' like *eat*, and that they are

1 Note that example (1a) is discussed in body paragraphs on pages 17-18 of Grimshaw, 1990). It is not stated as a distinct example (cf. Grimshaw 1990, examples 27 and 28).

internally heterogeneous.[2] There is also a significant body of literature focussing on languages which display different case marking when discussing some kinds of experiences when compared to non-experiential constructions. For example in Tsez, a Daghestanian language of the Northeastern Caucasus, dative case marking appears on the more agent-like argument of *ikʷay* "see" in (4), but not on *ac'* "eat" in (5) (Comrie & van den Berg, 2006, p. 137).

(4) *Tsez (Tsezic; ido1241)*

aħo-**r**	meši	b-**ikʷay**-si
shepherd-**DAT**	calf.ABS	III-**see**-PSTWIT

"The shepherd saw the calf"

(5) *Tsez (Tsezic; ido1241)*

žek'-**ā**	bišʷa	r-**ac'**-xo
man-**ERG**	food.ABS	IV-**eat**-PRS

"The man eats the food"

In each of these cases, the entire picture is more complex than can be discussed here. For example, while in Tsez the more agent-like arguments of the verb for *see*, *know*, and *hear* are marked with the dative as in (5), the more agent-like arguments of the verb for

2 There are good reasons to distinguish *eat* from prototypical action verbs (cf. Newman, 2009), but I set this aside here on the assumption that *eat* can be differentiated from psychological predicates.

understand is not (*ibid*, p.131). For our purposes, the important aspect of these examples is that there is a significant body of research identifying and discussing the ways languages differentiate experiential constructions, both from one another and from non-experiential constructions.

2.1.2 Overview of experiential constructions

As mentioned above, it is not the case that every language treats all "experiences" differently. The situation is summarised well by Dik (1981, p. 42-43 in Bugenhagen, 1990, p. 185):

> … It turns out that both within and across languages, these states of affairs, which we shall call "experiences", are expressed in different ways. It is as if languages cannot make up their minds as to how to express such experiences ... My impression is that it would be an oversimplification to assume that there is just one underlying structure for experiences ... Rather I believe that the various expression types found in natural languages reflect different conceptualizations of the notion of experience.

Researchers who are interested in investigating the variety of ways languages differentiate experiences are therefore left with the problem of defining a class of constructions which are used in the functional domain of experience so as to capture all of this variety. Studies

interested in this problem can be broadly grouped together as those studies which focus on “experiential constructions” in the sense that they first define the functional domain of experience before considering the constructions which appear in that domain for a language or group of languages.

For example, one set of studies focus on comparing experiencer encoding across African languages (e.g. Reh and Frühwald, 1996; Reh, 1998; Simon and Reh 1999; *inter alia*; cf. Becher, 2003; Zoch, 2020). Another group of studies apply the Natural Semantic Metalanguage approach to understand the semantic aspects of grammatical functions associated with experiencers, notably in Ameka (1990) on Ewe and Bugenhagen (1990) on Mangap-Mbula. A third group of studies builds on concepts derived from the UNITYP project (cf. Fedriani, 2014; Luraghi, 2021; Verhoeven, 2007; *inter alia*).

This subsection introduces some key notions which recur in this research, and which more clearly define the notion of the experiential construction. I largely focus on Verhoeven’s (2007) characterisation of experiential constructions, which appears to be the most parsimonious and comprehensive account of the topic and which aims to synthesise key aspects of earlier research. The scope of this review is also suitable for an Honours thesis, given a detailed overview of the similarities and differences between these research traditions is not possible here.

I take the experiential situation as a starting point to get to an understanding of experiential constructions.[3] This begins with the more general notion of the *situation*. A *situation* is defined as a kind of cognitive representation composed of two parts: *situation members* and a *situation core* (Lehmann et al., 2000). Situations can be categorised into different functional domains, which are "the universally relevant communicative and cognitive functions that language is assumed to ultimately serve." (Verhoeven 2007, p.4). As a generalisation, the functional domain of experience can be understood as containing "those concepts that are related to the consciousness of human beings and higher animals. More specifically, bodily feelings and sensations, feelings of emotions, concepts related to cognition, volition, and perception" (Verhoeven 2007, p.59; cf. Verhoeven 2007, Chapter 3).[4] The *experiential situation* then, is the situation associated with the functional domain of experience. I return to the various subdomains of the functional domain of experience below.

Considering now the component parts of the experiential situation, Simon & Reh (1999) and Zoch (2020) identify two situation members for experiential situations: the experiencer and the stimulus. (Verhoeven,

3 For greater detail on the concepts underpinning *experiential constructions* as introduced here, see especially Chapters 2 and 3 of Verhoeven (2007), and references therein.

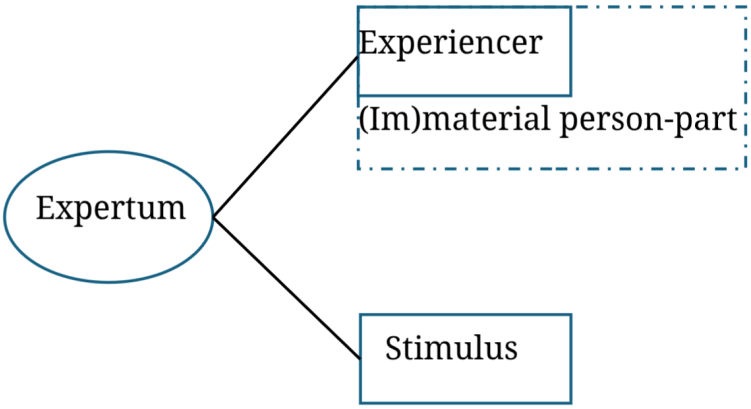
4 The question of cultural-specificity as opposed to universality of the functional domain of experience is briefly returned to with respect to psycho-collocation in Chapter 6.2 (cf. Enfield 2002; Verhoeven 2007, p.39-41).

2007) develops this picture further by also including (material or immaterial) person-parts (e.g. 'heart', 'mind' etc.) as a kind of second-order situation member given person-parts are involved in the situation only to the extent that they are parts of the experiencer (*ibid*:52).[5] Figure 2.1 visualises these different components of the basic experiential situation. In Figure 2.1., the *expertum* refers to the more general notion of the "situation core" mentioned above as it relates to an experiential situation. The situation core is generally expressed as a predicate, and is understood as constituting the experience itself: the sensation, cognition, perception etc. (Lehmann et al., 2000; Verhoeven, 2007).[6]

5 The term *person-part* (as opposed to *body-part* etc.) will be used here given both material and immaterial parts of a person are relevant for understanding psycho-collocation. Because material parts of the person are most often associated with the word body, it is inappropriate to simply use the term body-part here. Instead, *person-part* will be used, given the term encompasses immaterial parts and has been defined as the "unique beginner" for studies relating to the body elsewhere (cf. discussion in relation to semantic change in Wilkins 1996).

6 Simon and Reh (1999) and other studies in the same tradition (e.g. Reh and Frühwald, 1996; Reh, 1998; Becher 2003; *inter alia*) use the terms *experitum* rather than *expertum*, and *phenomenon* rather than *stimulus*. Later studies following Verhoeven (2007) use the terms *expertum* and *stimulus* for the same concepts, as I have chosen to do here.

Figure 2.1. Components of a basic experiential situation (per Verhoeven 2007, p.52



The situation members for an experiential situation - the experiencer, person-part (of the experiencer), and stimulus - can be defined broadly as in Table 2.1. Note that, similar to semantic roles, these ought to be understood as a useful generalisation rather than necessary and sufficient conditions (cf. Andrews, 2007, p.137, 140). Given an experiential situation, an experiential construction is the linguistic means of construing that situation.[7] That is, experiential constructions are the language used to talk about experiential situations. Psycho-collocation, the focus of this thesis, can be understood as one kind of experiential construction. By focussing on the kind of structures available in a language that express experiential situations, this thesis

7 *Construal* in this sentence is meant in the sense used in cognitive linguistics, where in the broadest sense construal can be understood as the conceptualisation of a given situation (Croft & Cruse, 2004, Chapter 3; Luraghi, 2021, p.3-6). This is important for comparing the different kinds of experiential constructions used within a language but can be put aside here.

broadly adopts an onomasiological approach, which will be introduced in Subsection 3.1.

Table 2.1. Members of the basic experiential situation (per Verhoeven 2007, p.52)

Situation Member	***General Definition***
Experiencer	"The experiencer participates through its physicalness and intellectuality in the situation which may be linguistically rendered by the use of material or immaterial body or person-part nouns."
∟*(Person-part)*	"… involved in the situation as parts of the experiencer. They are thus thought of as being on a secondary level; their participation being given by the experiencer … The experiencer is sympathetically affected when its parts are affected."
Stimulus	"A cover term for the situation member that triggers, causes, initiates the experiential situation, or to which the experience is directed towards."

There are a number of aspects of the relationship between these parts of the experiential situation that researchers have understood as important for experiential constructions. These include control, affectedness and volition, among others (cf. e.g. Malchukov, 2015; Næss, 2007, Chapter 8; Malchukov 2015; *inter alia*). There are two notions which are worth drawing out here: (a) the sympathetic relationship between the person-

part and the experiencer, and (b) the “relationality” of the situation core. They are discussed in turn below.

The sympathetic relation is understood as the “relation such that whatever affects the part necessarily affects the whole, and what affects the whole necessarily affects one of its parts.” (Lehmann, 2022, p.20). This is significant insofar as it differentiates the relationship between the person-part and the experiencer from other part-whole relations.

The “relationality” of the situation core refers to the notion that some situations require a semantic relationship between their elements (cf. Lehmann, 1985, p. 70-77). A significant corollaries of the relationality of the situation core for our purposes is the valency of the verb (cf. Verhoeven 2007, p.20). Note that both the sympathetic relationship and the relationality of the situation core are kinds of a more general notion of relationality. I return to these concepts with greater detail in Chapters 5 and 6.

2.1.3 Subdomains of the functional domain of experience

As a way of understanding the effect of different kinds of experiential situation, researchers divide the functional domain of experience into subdomains. There are several different ways of dividing things, though it appears that the various approaches for doing so all stem from earlier work by Tsunoda (1981, 1985) on case alternation. As such, there is a broad similarity between them. Taking three studies as an example,

Becher (2003; following Reh, 1998) distinguishes six subdomains, Verhoeven (2007) distinguishes five, and Malchukov (2015) distinguishes four in one part of a larger semantic map. These are given in Figure 2.2.

Figure 2.2. Subdomains in Becher (2003), Verhoeven (2007), and Malchukov (2015)

Becher (2003)	Perception	Cognition	Conception	Volition	Emotion	Physical sensation
Verhoeven (2007)	Perception	Cognition		Volition	Emotion	Bodily sensation
Malchukov (2015)	Perception/Cognition			Emotion		Sensation

Broadly, Verhoeven describes the ordering of subdomains from left to right in the figure above in terms of (increasing) dynamicity and (decreasing) relationality (Verhoeven 2007, p.352). Malchukov (2015) describes the ordering of subdomains above in terms of functional constraints such as whether a construction encoding a semantic role or assimilate to the typical pattern associated with bivalent verbs. In the remaining chapters of this thesis, I follow Verhoeven's (2007) conception of the experiential domain of experience's subdomains. I summarise these subdomains in Table 2.2 following Verhoeven (2007, p.41-51). There are several caveats that come with these subdomains. Most notably, it would be a mistake to assume these subdomains are entirely discrete. In these cases, a situation can be understood as part of more

than one subdomain. For example, a situation where an experiencer feels envious might be placed into the emotion and volition subdomains (Verhoeven 2007, p.42).

Table 2.2. Overview of the functional domain of experience's subdomains following Verhoeven (2007)

Subdomain name	***Description***
Bodily sensation	This subdomain is "concerned with feelings related to the experiencer's body" (Verhoeven 2007:42). Constructions associated with this subdomain are "prototypically states, but more dynamic situations are also included such as changes of state or processes" (*ibid*:43). ***Examples:*** **feel cold, feel sick**
Emotion	This subdomain includes "cognitively based feelings" (Verhoeven 2007:44; following Wierzbicka 1999). They are "biologically founded, interpersonal and have a social basis" (Verhoeven 2007:44). The emotion subdomain is associated with a causal chain with three events: emotional state, psychological reaction and/or other behaviours (*ibid*; following Radden 1998:273). It is "the most varied among the experiential subdomains from a semantic as well as grammatical point of view and has therefore been the motivation of most studies" (Verhoeven 2007:44). ***Examples:*** **feel fear, feel happiness**
Volition	This subdomain is "understood here as being comprised of two main areas of meaning, i.e., psycho-physical meanings… and meanings connected to intention" (V2007:48). The first is related to notions associated with liking, wishing and desiring, the second with intending and planning. ***Examples:*** **want, need**
Perception	This subdomain "includes those concepts that refer to the intake of information through the senses" (V2007:49). These can be either stative or dynamic (contrast see and look), though stativity is prototypical "since such situations are linguistically prone to having a special experiential coding" (V2007:50) ***Examples:*** **see, hear**
Cognition	This subdomain includes "internal experiences based on mental functions" (V2007:47). For example, situations that involve the presence or absence of information, which involve conceptual activity, or which involve attitudes towards the truth of propositions. ***Examples:*** **know, think**

2.2 Psycho-collocations

Having framed psycho-collocation as a kind of experiential construction, I now focus on psycho-collocation in its own right. In Section 2.2.1 I introduce psycho-collocation, and note that studies of psycho-collocation in the Mainland Southeast Asian language area (MSEA) restrict the term as presented in Matisoff (1986).[8] I then discuss four other areas of interest for this thesis from the study of psycho-collocation in MSEA. Section 2.2.2 expands on the semantic domains where psycho-collocations are used. I then note research which transplants psycho-collocation outside of MSEA (in Section 2.2.3). There are other aspects of psycho-collocation that are significant but which I defer for the time being. For example, I discuss psycho-collocations as idioms in Section 2.3 and return to the relationship between the participants in the experiential situation (per Verhoeven 2007) and the components of the psycho-collocation in Section 6.1. As a matter of scope, I focus on psycho-collocations where psycho-nouns are person-parts in this thesis, though I briefly return to the topic in Section 6.1.

8 It is important to note that this thesis does not investigate every aspect of psycho-collocation identified by studies on the topic. With this in mind, in this section I only focus on the aspects of psycho-collocation which appear to be both central to the concept and amenable to comparison with languages beyond MSEA. This is in line with my broader aim in this thesis of refining the notion of psycho-collocation for application beyond MSEA. With this is mind, this should not be seen as exhaustive. Some additional aspects of psycho-collocation are given by Vittrant (2013, p.260).

2.2.1 Broad and narrow psycho-collocation

To understand the application of *psycho-collocation* beyond MSEA, it is important to first discuss how the concept has been defined within MSEA. In doing so, it becomes clear that there is a difference between *psycho-collocation* as the term is used in Matisoff's (1986) and as it is used in later studies.[9] Let's first return to the definition of psycho-collocation in Matisoff (1986) which was given in Chapter 1. The definition is reproduced in full below.

> "a polymorphemic expression referring as a whole to a mental process, quality or state, one of whose constituents is a *psycho-noun*, i.e. a noun with explicit psychological reference (translatable by English words like HEART, MIND, SPIRIT, SOUL, TEMPER, NATURE, DISPOSITION, MOOD).

9 It should be noted that this is by no means a failing of research on psycho-collocation since Matisoff (1986) given shifts in the use of linguistic terminology over time are common in linguistics. A well-known example of this are terms for parts of speech, which derive from Greek and later Latin grammatical traditions (cf. Robinson 1997, Chapters 2 and 3 on parts of speech). Considering this context, differences in the use of *psycho-collocation* can be seen as a reflection of a more general tendency in linguistics. The change in the use of *psycho-collocation* appears to be more minor and subtle than cases like this but nonetheless should be considered for the sake of talking about psycho-collocation beyond MSEA.

> The rest of the ps[ycho]-collocation[10] constrains morphemes (usually action verbs or adjectives) that complete the meaning. This element we call the *psycho-mate*. We may then speak of *psychomatic verbs* and *psychomatic adjectives* – i.e. verbs or adjectives that combine in tight collocations with psycho-nouns. Thus:
>
> Psycho-noun + Psycho-mate = Ps[ycho]-collocation"
>
> (Matisoff 1986, p.4; emphasis and capitalisation in original).

Studies of psycho-collocation in individual MSEA languages consistently refer to some form of this definition when defining psycho-collocation. This is the case, for example, in Jaissier's (1990, p.160-161) study of White Hmong, Oey's (1990, p.142) study of Malay, VanBik's (1998, p. 201) study of Hakha Lai, and Vittrant's (2013, p.260) study of Burmese, among others. However, while these studies all refer to the same source, they also clarify some aspects of Matisoff's definition. Some of these are overviewed in Table 2.3, with indicative (but non-exhaustive) references.

10 Matisoff (1986) uses *psi-collocation* as a shortened form for psycho-collocation. I use psycho-collocation here for consistency with later literature.

Table 2.3. Changes in the use of the term psycho-collocation since Matisoff (1986)[11]

Aspect of psycho-collocation in Matisoff (1986, p. 4)	**Change in later studies**	**Reference(s)**
Psycho-collocations are polymorphemic	Typically have two monomorphemic parts: the psycho-noun and the psycho-mate	Mulder (2021); Vittrant & Watkins (2019, p.662)
A psycho-noun is any noun with "explicit psychological reference"	A psycho-noun is a typically a person-part noun	Enfield (2021, p.222); Oey (1990, p.142); Van-Bik (1998, p.201); Vittrant (2013, p.260); Vittrant & Watkins (2019, p.662); *inter alia*
A psycho-mate is all other (non-psycho-noun) morphemes which "complete the meaning", usually action verbs or adjectives	A psycho-mate is typically a (stative) verb	Van Bik (2023, p.102); Vittrant & Watkins (2019, p.662); *inter alia*

11 Table 2.3. should be interpreted as a general statement about how more recent studies restrict the notion of *psycho-collocation*. Studies of psycho-collocation in MSEA differ from Matisoff (1986) to different degrees, though they appear to agree that a psycho-noun is a person-part. With respect to the other two restrictions studies vary more widely. For example, Oey (1990) discusses psycho-collocations with multiple morphemes, and Vittrant and Watkins (2019, p.662) are more specific in their definition by defining a typical psycho-mate as a stative verb.

It appears that for the majority of authors writing about psycho-collocation in MSEA, aspects of the restricted usage of psycho-collocation given in Table 2.3 are taken to be the most common, but not exclusive, kind of psycho-collocation. These should not as significant divergences but instead specifications about psycho-collocation which come out as a matter of scoping an article's topic. This is illustrated by Vittrant and Watkins' guidelines for describing Mainland Southeast Asian languages, which note that psycho-nouns "*typically* pair a part of the body with a stative verb to express an emotion or feeling" (2019, p.662; emphasis mine). It follows that *atypical* psycho-collocation diverges from this model in some way. An example from White Hmong is (6) below, which differs from typical psycho-collocation in that it includes two person-parts: *siab* 'liver' and *ntshav* 'blood'.

(6) *White Hmong (Hmong-Mien; hmon1333)*
siab txia ntshav
liver ooze blood
"wicked, evil, cruel" (Jassier 1990, p.166)

A focus on 'typical' psycho-collocation acknowledges that there are other psycho-collocations which are part of the category, but which are not being investigated. Some of these restrictions are likely the result of the typological profile of languages in the MSEA language area. For example, some authors describe the psycho-mate as necessarily being a kind of verb, as opposed to considering psycho-mates which are other parts of

speech (e.g. Van Bik 2023, p.102). This is plausibly because words which are typically analysed as adjectives outside of MSEA tend to be analysed as either verbs or verb-like for languages in the MSEA language area (cf. Enfield 2021, p.211-214). Similarly, Verhoeven (2007, p.95) notes that there are constructions in African languages which are similar to psycho-collocations, but which notably tend to have abstract nouns rather than person-parts as the psycho-noun. This all leads to an important point for our purposes: because the restricted scope of *psycho-collocation* in languages in the MSEA language area may be motivated by the typological profile of that area, one cannot assume that 'typical' psycho-collocation in languages outside of MSEA is restricted in the same ways from the outset.

Given Matisoff originally intended psycho-collocation to apply across languages, as is evident in his application of the term to some Standard Average European languages, it can reasonably be assumed that it is possible to apply the notion to languages outside of MSEA (cf. Matisoff 1986, p.3-4 and examples throughout; cf. §2.2.3). A key part of this thesis is investigating whether that is an appropriate assumption (cf. Question 1; Chapter 1). In doing so, it is worth noting that more recent applications of *psycho-collocation* in MSEA tend to consider a subset of the more general notion in Matisoff (1986). As an attempt to differentiate original and more recent conceptions of psycho-collocation, I refer to the

restricted use of *psycho-collocation* in Table 2.2 as “narrow psycho-collocation”, given more recent uses of *psycho-collocation* appear to refer to a subset of the phenomena Matisoff (1986) describes using the term. Psycho-collocation as it appears in Matisoff (1986) might then be called “broad psycho-collocation”. I attempt to clarify the difference between broad and narrow psycho-collocation with reference to phraseology in Section 2.3.

2.2.2 Semantic domains of psycho-collocation

Considering experiential subdomains discussed in Section 2.1.2, it is worth briefly noting similarities and differences between those subdomains and the kinds of meanings expressed with psycho-collocations. Table 2.4 collects some of the different semantic domains attributed to psycho-collocations in studies focussing on MSEA.

Table 2.4. Differences in the semantic range of psycho-collocation for some publications on MSEA languages

Semantic field/domain associated with psycho-collocation	***Reference***
Divides into: A) "Intellection and attitudes towards experience", which is further divided into: i) intellection ii) volition iii) experiential attitudes B) "Qualities of character/ traits of personality… qualities that are felt to be quasi-permanent, not subject to change from moment to moment" C) "Emotions, feelings, moods … express feelings moods, or states of mind (cf. French *états d'espirit*) which are temporary or momentary, evoked by a particular mental or physical event."	Matisoff (1986, p.5-6)
Psycho-collocations are divided into "moral characteristics/personality traits", "mental processes/states", "anger", "satisfaction", "fear", "confusion" and "other emotional states"	Jaissier (1990)
"expresses an explicit psychological state"	Van Bik (2010, p.141)
"make explicit reference to body parts or organs as the locus of emotions or psychological states"	Vittrant (2013, p.260)
"express an emotion or feeling"	Vittrant and Watkins (2019, p.662)
"denote states and properties in the interconnected domains of cognition, emotion, and personality: concepts like happy, sad, surprised, frightened, impulsive, kind, and arrogant."	Enfield (2021, p.222)
"emotions, qualities of personality, and intellectual activities"	Mulder (2021, p.1213)

White Hmong (Hmong-Mien; hmon1333)

If we accept that these authors are all referring to what is largely the same set of meanings, we can talk about differences between this field and the subdomains of the experiential situation noted in Section 2.1.2. There are three cases that should be noted here: qualities, perception and bodily sensation.[12] What Matisoff (1986:5) refers to as "qualities that are felt to be quasi-permanent, not subject to change from moment to moment" appear to be equivalent to "property concepts" in Verhoeven's (2007) analysis. For Verhoeven, the term *property concept* refers to a situation's "internal temporal structure" – a grouping of notions from theories of aspect like dynamicity (stative vs. dynamic), telicity (atelic vs. telic) and durativity (durative vs. punctual) (cf. Verhoeven 2007:15-18 and references therein). Considering this, properties can be understood as orthogonal to the subdomains given in 2.1.2. to the extent that they describe how the situation extends over time as seen in Matisoff's use of "quasi-permanent" (cf. Verhoeven 2007:46).[13] Following from this, one might talk about an emotional property concept as in (7), or a non-

12 There are some other small things which I put aside here. For example, Matisoff (1986, p.5) groups "volition" alongside "intellection and attitudes towards experience", where in other relevant publications volitional meanings like *WANT* appear to be understood as emotions (e.g. in Bugenhagen 1990).

13 Matisoff (1986, p.6) also notes that the dimension of "subjectivity/intersubjectivity" crosscuts the classification of psycho-collocations he gives which appears in Table 2.3. I return to this in Chapter 5.

experiential property concept as in (8), which is described as a moral property in Jassier (1990).

White Hmong (Hmong-Mien; hmon1333)

(7)

siab kub

liver hot

“Quick, hot-tempered, prone to anger”

Jaissier (1990, p.176)

(8)

siab ncaj

liver straight

“Morally honest and upright”

Jaissier (1990, p.164)

There are two experiential subdomains given in 2.1.2 that appear to go beyond the domains given in Table 2.3.: perception and bodily sensation. Psycho-collocations which can be understood as part of the bodily sensation subdomains can be found on research on psycho-collocation in MSEA. This is demonstrated for bodily sensation in (9). The perception subdomain appears to be absent in accounts of psycho-collocation in MSEA, a point I come back to in Section 5.1. Keeping this in mind, I tentatively assume that psycho-collocations can be part of the perception and bodily sensation subdomains for the purpose of data collection (cf. §3.2.). The extent to which the descriptions given in Table 2.3 or the subdomains given by Verhoeven (2007) capture the meanings expressed with psycho-collocation will be further explored in Section 5.3.

(9) *Hakha Lai (Tibeto-Burman; haka1240)*

Ka	luŋ	ʔa	mit
1SG.POSS	heart	3SG.SUBJ	extinguish.I

“I feel dizzy” (lit. “My heart is extinguished [like a light or a fire]”)

(Van Bik 1998, p.213)

2.2.3 Psycho-collocation outside of Mainland Southeast Asia

In investigating relevant research on psycho-collocation beyond Mainland Southeast Asia, it quickly becomes clear that there are various bodies of research discussing closely related phenomena using different terminology. I’ve given some of these terms in Table 2.5 as an indication. Note that in some cases can be understood as a kind of psycho-collocation (e.g. complex emotion predicates in Musgrave, (2006)) and others as larger categories containing psycho-collocation (e.g. *experiencer-as-possessor* constructions in Bickel (1997, 2004)).

Table 2.5. Terms used outside MSEA which appear closely related to psycho-collocation

Term	*Definition*	*Reference*
Bodily expressions	"related feelings and the parts of the body where they are thought to be located and the kind of activity or state associated with the body part(s)."	Ameka (2002, p. 29)
Body image idioms[14]	"those [expressions] which consist of primarily the terms for the body parts plus predications about them"	McElhanon (1977, p.3)
Metaphorical body part expression	"a body part phrasal which usually renders the emotional situation as the physical quality of a body part and encodes the experiencer-participant as the possessive attribute of the respective body part term."	Becher (2003, p,14)
Complex emotion predicates	"complex predicates consist[ing] of a body-part noun, which is often possessed, and an attributive word denoting the state of the body-part noun"	Musgrave (2006, p.228)
Experiencer-as-possessor constructions[15]	"experiential expressions … [which] have the experiencer marked as the possessor of an affected body part or feeling."	Bickel (2004, p.86)
Body part-centred constructions	"In this construction a body part is expressed as nominative Noun Phrase, which metonymically refers to the Experiencer"	Fedriani (2014, p.47)

The notion of *psycho-collocation* appears to be a reasonable starting point for considering the similarities and differences between the various phenomena identified by these researchers. While the scope of

14 cf. "body image expressions" in McElhannon (1975) and applications in Bugehagen (1990, 1995, 2001). Note also related analysis in McElhannon (1992).

15 cf. the earlier "possessive of experience" in Belhare (Bickel 1997).

an Honours thesis does not allow the space to compare the merit of all this alternate terminology, I believe taking psycho-collocation as the point of comparison is appropriate given existing applications of *psycho-collocation* both within and beyond MSEA.[16] Notable studies that explicitly apply psycho-collocation outside of MSEA include Bickel (1997) on Belhare, Verhoeven (2007) on Yucatec Maya, Von Prince (2017) on Daakaka and Weissmann (2023) on Teposcolula Mixtec, and Lahaussois (In Press) on Thulung, among others. Given the range of relevant literature outside of MSEA is vast, I do not aim to overview it here and instead give references to appropriate literature outside of MSEA where appropriate in the sections that follow.

2.3 Phraseology and idioms

As noted in Chapter 1, a key focus of this thesis is investigating the kind of category psycho-collocation is, particularly with respect to its application beyond the Mainland Southeast Asian language area. This section notes existing characterisations of psycho-collocations as idioms and/or collocations and clarifies this terminology with respect to Mel'čuk's (2023) *General Phraseology*.[17] A significant import of this

16 Though note again the variations between conceptions of psycho-collocation as originally conceived by Matisoff (1986) and in later studies; cf. 2.2.1.

17 That I follow Melčuk's (2023) characterisation of phraseology here should not be taken as denying the value of research on idioms in other research traditions. While psycho-collocations and related phenomena might be productively studied in greater detail under these frameworks, I have chosen to follow Melčuk's (2023) general phraseology here primarily because of its brevity

approach is the notion of compositionality, which I introduce here. The concepts introduced in this section serve as both a way of clarifying existing terminology under a single framework, and as background for the analysis of psycho-collocations in Section 5.1. Considering Section 2.2.1, I conclude the chapter by giving two definitions of psycho-collocation with respect to General Phraseology: one that reflects Matisoff's (1986) original conception (which I label "broad psycho-collocation"), and another which agrees with more recent changes in the use of the term (which I label "narrow psycho-collocation").

2.3.1 Psycho-collocations as either idioms or collocations

The notion of psycho-collocation as a kind of idiom plays no apparent role in Matisoff's (1986) original study. Nonetheless, idiom has played a role as a general label for psycho-collocations in later research. It is not the case that all research in the area applies the term in the same way. For example, Mulder (2021) understands psycho-collocations as idioms by definition, while Van Bik (1998) distinguishes psycho-collocations which are idioms from those which are not. Given authors since Matisoff (1986) appear to agree that the notion of idiom is relevant, but disagree on the term's application, it is therefore important to clarify the notion of *idiom* for the purpose of comparison. In doing so, I follow Melčuk's (2023) general phraseology.

and precision.

At the broadest level, Melčuk (2023) is interested in *phrasemes*, which are defined as a complex linguistic sign where "the selection of one of its constituent elements by the speaker is constrained" (*ibid*, p.240).[18] Psycho-collocations are understood as constrained to the extent that the meaning of the psycho-collocation cannot be achieved with only one of either the psycho-noun or psycho-mate. Melčuk (2023) subdivides the notion of phraseme several times. One subdivision is the *idiom*, defined as a phraseme which is non-compositional, and where all components of the phraseme are non-grammatical lexemes (*ibid*, p.241).[19, 20] The first part of this definition, relating to non-compositionality, is most relevant for our purposes. Here it is sufficient to understand non-compositionality as the inability to represent the meaning of one linguistic sign as the regular "sum" of more than one other linguistic signs (cf. Melčuk 2023, p.248). Somewhat confusingly, Melčuk (2023)

18 Strictly, it must also be the case that "if the constrained selection concerns only one sign within the phraseme, this sign is not a *grammatical* lexeme" (Melčuk 2023:24). This point is not relevant for the discussion of psycho-collocations here and so is noted but put aside.

19 i.e. it is a "lexemic phraseme" (cf. Melčuk 2023, p.40, 240). Lexeme is defined in Melčuk (2023) as "A one word expression", which includes the "set of wordforms and analytical-form phrases that differ only by inflectional significations", meaning *I* and *me* are the same lexeme, for example (Melčuk 2023, p.255).

20 Note that definitions in Melčuk (2023) have corresponding formalisations which are not reproduced here.

contrasts idioms with *collocations*.[21] The only difference between collocations and idioms for Melčuk (2023) is their compositionality, where collocations are compositional and idioms are not. These two classes can then be further subdivided. These divisions, first with respect to constrained selection and then with respect to compositionality, are relevant insofar as they give comparatively concrete criteria for distinguishing and subdividing the set of psycho-collocations. I return to the notion of the phraseme in Chapter 5. In Section 5.1.1 I outline an argument for psycho-collocations as demonstrating constrained selection, and thereby being phrasemes. In Section 5.1.4 I discuss compositionality, and conclude that while there is some evidence for the view that psycho-collocations can be either idioms or collocations, that evidence is insufficient to provide a definitive answer.

21 Considering the use of "collocation" by Melčuk, I should note that I chose to keep the term "psycho-collocation" here for several reasons. Firstly, I believe that linguistic terminology should preserve continuity to the extent possible (cf. Haspelmath 2025, p.36). Further, psycho-collocation appears (as far as I am aware) to be the most widely used term for the phenomena I discuss in this thesis. Finally, the word "collocation" is often understood in the sense used in corpus linguistics, where "collocation" refers to any co-occurring set of two or more lexemes (Barth & Schnell, 2021, p.14; cf. Vittrant 2013, p,260). Note that this sense of the term is closer (but not equivalent) to Melčuk's notion of the "phraseme", which in Melčuk's general phraseology includes all idioms and collocation (as defined by Melčuk).

2.3.2 Preliminary definitions of broad and narrow psycho-collocation

Based on the information given in the previous sections, in this subsection I tentatively propose a definition of (broad) psycho-collocation (cf. §2.2.1). This is useful as it allows me to investigate constructions which might be thought of as psycho-collocations but which are not labelled as such. While the following definition is preliminary, it has therefore structured the data collection and analysis given in the remaining chapters. I return to this definition in Section 6.1.

I first define broad psycho-collocation. (Broad) psycho-collocations are understood here as lexemic phrasemes whose:

1) Lexemes include at least one psycho-noun and one psycho-mate, and which
2) Denotes an experiential situation.

This definition can then be decomposed into a set of statements about what kind of expression psycho-collocations are.

(Broad) psycho-collocations are:

1. PSYCHO-COLLOCATIONS ARE PHRASEMES; cf. §2.3.1

complex linguistic signs where “the selection of one of its constituent elements by the speaker is constrained” (Melčuk 2023, p.240)

2. PSYCHO-COLLOCATIONS ARE MADE UP OF LEXEMES, cf. §2.3.1

This phraseme is only made up of lexemes: a “one word expressions”, which includes the “set of wordforms and analytical-form phrases that differ only by inflectional significations”, meaning I and me are the same lexeme (Melčuk 2023, p.255).[22]

3. PSYCHO-COLLOCATIONS INCLUDE AT LEAST ONE PSYCHO-NOUN AND AT LEAST ONE PSYCHO-MATE cf. §2.2.1

The lexemes which make up the psycho-collocation are divided into psycho-nouns and psycho-mates, where

- a psycho-noun is “a noun with explicit psychological reference (translatably by English words like HEART, MIND, SPIRIT, SOUL, TEMPER, NATURE, DISPOSITION, MOOD” (Matisoff 1986, p.4) and

22 As noted in subsection 2.3.1, statements (1) and (2) above make psycho-collocations “lexemic phrasemes” in Melčuk’s (2023) general phraseology,

- a psycho-mate are “morphemes (usually action verbs or adjectives) that complete the meaning [of the psycho-collocation]” (*ibid)*.

There must be, minimally, at least one psycho-noun and one psycho-mate in the psycho-collocation.[23]

4. PSYCHO-COLLOCATIONS ENCODE EXPERIENTIAL SITUATIONS cf. §2.1.2

Psycho-collocations encode an experiential situation. That is, a situation associated with a functional domain which “contains those concepts that are related to the consciousness of human beings and higher animals. More specifically, bodily feelings and sensations, feelings of emotions, concepts related to cognition, volition, and perception are included.” (Verhoeven 2007, p.39; cf.

23 This last sentence is not discussed in the preceding sections and is worth explaining further here. Psycho-collocations are defined here as requiring at least one psycho-noun and at least one psycho-mate. This sentence does two things in Matisoff’s (1986) definition. Firstly, it encompasses Matisoff’s (1986, p.4) notion that each psycho-collocation has *both* a psycho-noun and a psycho-mate (embedded in the idea of the psycho-mate “completing” the meaning and demonstrated by a lack of psycho-collocations made up of only psycho-nouns). The sentence also admits cases where there is more than one psycho-noun or psycho-mate. This is the case in four-part “elaborate expressions”, which are sometimes understood as including psycho-collocation in Matisoff (1986, p.35).

> ibid, p.4-5).[24] This means psycho-collocations are a kind of experiential construction.
>
> While they are considered peripheral in the study of experiential constructions more generally, qualities (which "refer to aspects of a person's NATURE, CHARACTER, DISPOSITION, TEMPERAMENT, or PERSONALITY – i.e. qualities that are felt to be quasi-permanent, not subject to change from moment to moment"(Matisoff 1986, p.5) are considered experiential constructions insofar as they conform to the larger class of psycho-collocations in a language.

Considering the above definition of broad psycho-collocation, we can reason about a "narrow" psycho-collocation, where tendencies that researchers have associated with typical (broad) psycho-collocation are taken as characteristics of all (narrow) psycho-collocations. For example, the body of research introduced in Section 2.2.1 suggests that:

---

24 It is assumed here that the semantic domains of psycho-collocation defined in Matisoff (1986) are always a kind of experiential situation. This is justified to the extent that the semantic domains mentioned in subsection 2.2.2 are understood as tendencies by Matisoff (1986, p.3). In comparing psycho-collocations across languages, I take it as important to not presuppose that the domains which are most apparent in MSEA are most apparent elsewhere. Note that Verhoeven (2007, p.93) justifies the use of the distinct term *experiential collocation* on similar grounds. Further, note that this aspect of the definition is understood as subsuming Matisoff's (1986, p.6) characterisation of psycho-collocations as requiring a "mental" meaning.

- psycho-nouns are most often person-parts, which includes physical/bodily parts of a person like “heart” and “eye” as well as abstract parts of a person like “mind”.
- psycho-collocations can typically be divided into two lexemes: a single lexeme categorised as a psycho-noun and a single lexeme categorised as a psycho-mate.
- Psycho-mates are typically verbs or adjectives
- The relationship between psycho-noun and psycho-mate is typically a kind of phrasal predicate, though they are also often analysed as forming a compound or having the psycho-noun incorporated into the psycho-mate.

Note that there are a number of other aspects of psycho-collocation which have been noted in the literature on psycho-collocations. I return to some of these in Chapters 5 and 6. In the sections that follow, I aim to consider broad psycho-collocation in five languages: Dalabon, Ewe, Hakha Lai, Mangap-Mbula, and Yucatec Maya. Considering the resources and the time available for completing an Honours thesis, here I only consider psycho-collocations where the psycho-noun is a person-part as a matter of scope.

3 Approach

This section aims to overview the approach taken in this thesis. In Section 3.1 I begin by broadly framing my approach in relation to onomasiological and (functional-)typological studies. In Section 3.2 I then introduce the set of languages considered in this thesis. Section 3.3 then focusses on the approach I took to constructing a database of psycho-collocations. While they are of relevance to this section, the methods for analysing this database are introduced as relevant in Chapter 5.

3.1 General characterisation of method and assumptions

With respect to data collection and interpretation, my approach was informed by a view common to (functional-)typological approaches: that variation is the normal (within and between) state of natural languages and that generalisations across languages may be used to understand that variation (cf. Croft 2002, Chapter 9). In practice, this meant that the approach taken here aims to compare psycho-collocation in five languages as a way of accounting for variation in the use of psycho-collocation as an experiential construction. This is done as a way of addressing the research questions given in Chapter 1. While these claims might then be used to make further assertions about experiential

constructions as a whole I only address those potential contributions in general terms (cf. Croft 2002, p.284; cf. *ibid,* Chapter 6).

Studies of experiential constructions such as those introduced in Section 2.1.2. are necessarily onomasiological (Carnesale, 2023; Fedriani, 2014, p. 17). This means that one first focusses on a way in which language is used before considering the ways languages instantiate that function (Kimoto et al., 2024, p.126; Mosel, 2006, p.84). This thesis builds on that onomasiological approach by considering a significant kind of experiential construction: psycho-collocation. As such, it is constrained by Matisoff's (1986, p.6) initial definition of psycho-collocations "referring as a whole to a mental process, quality, or state". Since there is not a consensus among linguists as to what counts as "mental", I tentatively broadened to the experiential situation as defined by Verhoeven (2007) as discussed in Section 2.1.2 and 2.2.2.. I return to the applicability of the notion that psycho-collocations are experiential constructions in Section 5.2.., given it quickly became apparent that there are psycho-collocations which go beyond what is considered a typical experiential construction in Verhoeven (2007).

Within the more general context of onomasiological approaches, this study adapts Evans' (2004, p.169) characterization of "the standard approach to semantically-based typology". While Evans' (2004) focus was

on experiencer objects in the Iwaidjan languages, the approach is applicable beyond those languages. It can be generalised as follows:

(a) Consider meanings that are typically associated with some construction (e.g. consider the kinds of verbs expressing meanings associated with non-nominative subjects in Russian, Hindi, etc.).

(b) Identify the construction type used to express those meanings in another language (or set of languages). (e.g. identify the construction types associated with the meanings considered in (a) for the Iwaidjan languages).

(c) Examine the semantic range of predicate lexemes used for this construction in these languages.

With respect to psycho-collocation, the concern of this thesis, Evans' approach is adapted though the following steps: **i)** Assume as a starting point Verhoeven's (2007) conception of the functional domain of experience, taken to contain the taxonomy of meanings associated with psycho-collocation in Matisoff (1986) (cf. §2.1.); **ii)**; identify psycho-collocation as a significant construction type for expressing those meanings in five languages (cf. §3.2.); **iii)** gather examples of psycho-collocation in these languages (following the definition of broad psycho-collocation given in Section 2.3.2.; cf. §3.3.); **iv)** examine the semantic range of these psycho-collocations, the semantics of their parts (the

psycho-noun and psycho-mate), and variation within the category with respect to their structure (cf. Chapters 4 and 5).

3.2 Languages considered

Five languages have been selected for the present study. They are (with language family and macro-area in brackets, following terms used by Hammarström et al., (2024)):

- Dalabon (Gunwinyguan, Australia)
- Ewe (Atlantic-Congo, Africa)
- Hakha Lai (Sino-Tibetan, Eurasia)
- Mangap Mbula (Austronesian, Papunesia)
- Yucatec Maya (Mayan, North America)

These five languages have been chosen out of careful consideration of several factors, particularly the aim of having as diverse of a language sample as possible while accounting for time constraints associated with an Honours thesis. Another important consideration is the general lack of data from studies which identify a construction conforming to psycho-collocation and the need to consider a range of person-parts. I briefly consider sampling below.

At the time of writing, psycho-collocation has only received dedicated study in a handful of languages, largely located in Mainland Southeast Asia (cf. e.g. Enfield 2021, p.222-227). This makes sampling a difficult

task, especially considering that this study intends to investigate the merit of extending the concept of psycho-collocation beyond the MSEA language. While this set of five languages cannot be said to be representative of the world's languages, it has been constructed following principles applied to larger samples such that it can form a useful starting point for future research into psycho-collocation and experiential constructions across languages. For example, the given languages are all from different phyla and macro-areas (following Hammarström et al. 2024), which are listed alongside the languages above. These areas are Australia (for Dalabon), Africa (for Ewe), Eurasia (for Hakha Lai), Papunesia (for Mangap Mbula) and North America (for Yucatec Maya). Appropriately representing the variety of languages within different phyla and macroarea is significant for larger samples (cf. Miestamo et al., 2016; Rijkhoff et al., 1993). Given the total number of languages sampled and time constraints, it is appropriate that I only consider distinct macro-areas and phyla rather than sampling within macro-areas or phyla. Finally, it should be noted that the languages under investigation differ in the amount and kind of relevant research available. I return to the implications of these limitations in Chapter 6.

3.3 Constructing a database of psycho-collocations

For the purpose of investigating psycho-collocation across languages, I constructed a database of approximately 700 examples across the five languages discussed in Section 3.2. I rely on this database of psycho-collocations for the comparison between languages given in Chapter 5.[25] Here I briefly outline the process of constructing this database and explain the sources of data for each language under investigation.

For the most part, authors identify a handful of person-parts which are taken as indicative. Where possible, I took this as the starting point for collecting examples of psycho-collocation. The notable exceptions to this were Ewe and Dalabon. For Ewe, I gathered psycho-collocations in general, then restricted that list to the most frequent person-parts (those with three or more unique noun-predicate pairings). For Dalabon, I collected examples from Ponsonnet (2014), which focusses on emotion language, then supplemented this list with examples from Evans (2007, 2021). This approach notably differs from comparative studies that take a restricted set of predicate meanings as their starting point (e.g. Evans, 2004; Hartmann et al., 2016)). Given I am interested in the kinds of semantic contribution of the person-part and the predicate, I return to possible implications of foregrounding the person-part in my approach

25 This database is given as a supplement.

in Chapter 6. Examples were collected from studies across the five languages under consideration as listed in Table 3.1. Note that other sources beyond those listed informed Chapters 4-6, but did not contribute to the database.

Table 3.1. Sources used to gather examples in the database

Language	***Sources used to gather examples in the database***
Dalabon	Evans, (2021); Ponsonnet, (2009, 2010, 2014b, 2014a, 2015)
Ewe	Ameka, (1990, 2002); Awuku, (2006); Sevor, (2015, Chapter 5)
Hakha Lai	Van-Bik, (1998); Yandt, (2019)
Mangap-Mbula	R. D. Bugenhagen (1990, 2001); S. E. Bugenhagen & Bugenhagen (2007)
Yucatec Maya	Verhoeven (2007)

The dictionaries given in Table 3.2 were used to gather more detailed information about psycho-nouns and psycho-mates. Where no dictionary entry could be found for a psycho-mate, I reproduced the gloss in the original source. Two dictionaries were necessary for Ewe given Westermann (1928) does not list part of speech. Examples of psycho-collocation were gathered from Bugenhagen and Bugenhagen (2007) but not other dictionaries given they were listed as sub-headwards underneath entries for person-parts. Issues relating to comparability and the reliance on bilingual English dictionaries are returned to in Chapter 6.

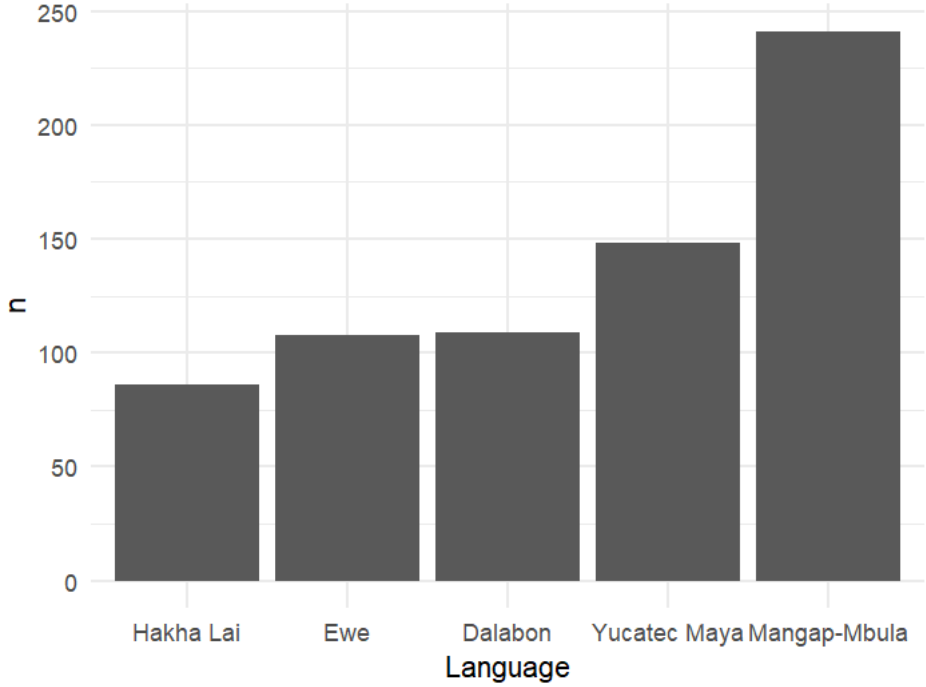
Table 3.2. Dictionaries used in database

Language	***Dictionaries used in database***
Dalabon	Evans et al., (2004)
Ewe	Rongier, (1995); Westermann, (1973)
Hakha Lai	D. Van Bik, (2014)
Mangap-Mbula	Bugenhagen and Bugenhagen (2007)
Yucatec Maya	Bricker et al. (1998)

In the database I include an example where that example denotes an experiential situation (cf. §2.1. and §2.3.2.), includes a person-part, includes a predicate, and is not a duplicate of another example. When determining whether an example was a duplicate, I excluded examples which only differed with respect to the referent of the experiencer or stimulus argument. Note that this also meant treating two examples which only vary in the number and person marking of participants as duplicate entries.[26] Overall, I collected ~800 examples. These distribute across languages per Figure 3.1.

26 This should not be taken as a claim that insights cannot be gained from considering these features (cf. e.g. Van Bik, 2010; Bickel, 2000).

Figure 3.1. Number of identified psycho-collocations by language



For each example in the database, I assigned an experiential subdomain. In many cases this information was given by the author. Where that is not present, I made a judgement about subdomain based on the gloss, and any additional comments by the author, with reference to the definitions in Verhoeven (2007). Additionally, I categorised the kind of relationship between the psycho-mate and psycho-noun as either an oblique phrase, as part of an argument, or as incorporated into psycho-mate. I return to other relationships, like phrasal verbs and cases where the person-part acts as the predicate, in Chapter 5. For each person-part and predicate, I collected a dictionary definition and part of speech

(including relevant information relating to word class etc.).[27] Additionally for psycho-nouns I collected information about whether the definition of that part includes only a corporeal meaning, both corporeal and incorporeal senses or only an incorporeal sense based on the dictionary definition. For psycho-mates I made a judgement about whether the definition could refer to an experiential situation, as introduced in Section 2.1 when it appears without a person-part based on the dictionary definition and gloss in the original example. Figure 3.2 reproduces the database entry for *ɖɔme nyo* 'to be generous/kind' as an example.

Figure 3.2. Example database entry

Index	***Language***	***Reference***	***Construction/ Example***	***Gloss***	***Psy-coll Definition***
125	Ewe	Awuku (2006:146)	ɖɔme nyo	stomach good	to be generous/kind

Person-Part	***Mate Gloss***	***Mate Definition***	***Mate PoS***	***Mate Experiential?***	***Subdomain***
dɔme 'stomach'	nyo 'good'	good, nice, fine, pretty; very, much (Westermann 1928:192)	Adjective (Rongier 1995:533)	N	EMO[28]

27 For the terms for word classes, I followed the relevant dictionary. I make no claim about the comparability of word classes across languages here. However, I do consider the significance of word classes within a given language in Chapters 5 and 6.

28 Where "EMO" refers to the emotion subdomain. Details of abbreviations used in the database are given as part of the supplementary materials for this thesis.

4 Overviews of Individual Languages

This chapter surveys key information about psycho-collocation in the five languages under consideration: Dalabon, Ewe, Hakha Lai, Mangap-Mbula, and Yucatec Mayae (cf. §3.2. on sampling). Each section below (§4.1-4.5) will consider a given language. Each subsection is divided into two parts. The first subsection includes information about the person-parts used in psycho-collocations, including their distribution across subdomains with respect to unique pairings of psycho-mate and psycho-noun. The second subsection overviews information about how those person-parts contribute to the formation of psycho-collocations more broadly. In Section 4.6 I give an overview of the contribution of this chapter and link back to the notion of broad psycho-collocation introduced in Section 2.2.1.

4.1 Dalabon

Dalabon is a Gunwinyguan language of South-western Arnhem Land in Australia. In Dalabon, emotion expressions feature a large number of person-part words. These person-parts most often appear in noun-incorporation constructions, similar to other neighbouring languages (Ponsonnet & Laginha, 2020; (Ponsonnet et al 2020:8-9; O'Keeffe et al 2020). This section primarily draws from Ponsonnet (2009, 2014a, 2014b, 2015), Evans (2007, 2021) and Evans et al. (2004). Much of Ponsonnet's

work (e.g. Ponsonnet 2014a, 2014b, 2015, 2017) primarily focuses on emotions in Dalabon. Evans (2007, 2021) and Ponsonnet (2009) include additional relevant information for cognition and perception predicates. Bodily sensations do not appear to be the central focus of any publication on Dalabon, meaning information here has been gleaned from research where bodily sensation is not a central topic.

4.1.1 Person-parts in Dalabon Psycho-collocations

In Dalabon, all parts and attributes of animate nouns[29] are part of a noun class that is semantically well-identified and has distinct morphosyntactic properties (Ponsonnet 2015, p.21-22). This includes cases where the person-part is incorporated into the verb complex and those where the person-part is not. Here I focus on a small sample of person-parts which appear in verbal compounds, returning to unincorporated person-parts in 4.1.2. It was necessary to adjust my method for sampling Dalabon psycho-collocations. I drew from the most frequent person parts which are included in Appendix 1 of Ponsonnet (2014), Additionally, I added *dje-no* 'nose, nostrils, face', *koh-no* 'eyes, gaze' and *men-no* 'consciousness' on the basis that lexemes with similar meanings are used in subdomains other than emotion in other

29 While Ponsonnet (2014a; 2015) uses the term "parts of animates", I will use person-parts going forward for the sake of consistency. It is assumed that all person-parts are in the class of parts of animates.

languages in the sample. Definitions for the sampled person-parts appear in Table 4.1.

Table 4.1. Dalabon person-parts in database

Person-part	***Definition***
dje-no	“nose/nostrils/face” (Ponsonnet 2015, p.405)
kodj-no	“head, either for a human being or for an animal” (Ponsonnet 2009, p.21)
dengu-no	"foot/toe" (Ponsonnet 2015, p.15)
kangu-no	“meaning (possibly among other things) the belly, the stomach, or ‘abdominal container’” (Ponsonnet 2010, p.373)
men-no	“an attribute of persons and animals that refers to their consciousness” (Ponsonnet 2009, p.21); “ideas, mind, consciousness judgement” (Ponsonnet 2015, p.408
ngerh-no	“heart, breath” (Ponsonnet 2014b, p.407)
NGURLK	Does not appear as an independent lexeme but “probably related to *ngurl-no* ‘heart’ (etymology accessible to speakers)” (Ponsonnet 2014b, p.407)
koh-no	"eyes, gaze" (Ponsonnet 2014b, p.408)
yolh-no	“an abstract attribute of the person: someone’s psychological and physical condition, energy, pep, enthusiasm, as well as motivations and desires (emotional), and intentions (intellectual).” (Ponsonnet 2014b, p.281)

With respect to the subdomains of the experiential situation discussed in Section 2.1.3., the examples I gathered distribute across those subdomains as shown in Table 4.2 (Note again that the emotion subdomain has massive internal complexity).

Table 4.2. Person-parts and subdomains in a sample of Dalabon Psycho-collocations (unique noun-mate pairings)

Subdomain / *Person-Part*	***Cognition***	***Perception***	***Volition***	***Emotion***	***Bodily Sensation***
dje-no 'nose, nostril, face'				✓✓	✓
kangu-no 'belly'	✓✓		✓	✓✓	✓✓
kodj-no 'head'	✓✓			✓✓	✓✓
koh-no 'eyes, gaze'		✓✓		✓✓	
men-no 'consciousness'	✓✓			✓	
ngerh-no 'heart, breath'			✓	✓✓	
NGURLK 'heart'			✓	✓✓	
yolh-no 'pep, feelings'	✓*		✓✓	✓✓	✓✓ **

✓ – At least one unique mate-noun pair.
✓✓ – Three or more unique mate-noun pairs.
*Knowing how someone feels
**Can be used for bodily sensation where the psycho-collocation was also coded as being in the emotion subdomain.

I will now consider the person-parts used for psycho-collocations with respect to each of the subdomains listed in Table 4.2. The **cognition** subdomain includes examples with *kodj-no* 'head', *kangu-no* 'belly', *men-no* 'consciousness' and *yolh-no* 'feelings' (cf. Evans 2007, 2021; Ponsonnet 2009). M*en-no* 'consciousness 'is an abstract attribute located in the head

(but without a particular body-part as its referent) which is used when discussing thoughts, views/judgements and intentions (Ponsonnet 2009, p.21). As such *men-no* is closely associated with the cognition subdomain, though it also occurs in compound with *yer* 'shy.away' to refer to emotional experiences translated as shame and embarrassment, which can be grouped with emotions (Ponsonnet 2014, p.155, 409). Additionally, *Kodj-no* 'head' is a visible part of the body understood as the container and processor of mental states, and which relates closely to the "socially defined subject" (Ponsonnet 2009, p.23). While *yolh-no* 'feelings' predominantly appears in the emotion subdomain, it also appears in the compound *yolh-nan* 'see what someone is feeling or thinking' (lit. 'pep/feelings+see') (Ponsonnet 2014, p.445).

Although predicates in the **perception** subdomain appear with person-parts, they appear to most often incorporate the stimulus (i.e. thing perceived; e.g. Ponsonnet 2015, ex.34), perhaps limited by the more general constraint that incorporated nouns in Dalabon must express the absolutive argument (Ponsonnet 2014a, p.64).[30] There are some cases of the perception subdomain in the database however, like *koh-no* 'eyes, gaze' in *ko-wurrh(mu)* 'be surprised by the sight of something' (lit. 'eyes/gaze+surprised, frightened'; Ponsonnet 2014b, p.429) and *koh-na*

30 Where the absolutive argument is either the sole argument of an intransitive clause or the more patient-like argument of a transitive clause.

‘keep an eye on’ (cf. Ponsonnet 2014b, p.213), which both occur in situations denoting an emotion with alongside behaviour related to sight/gaze. However, for the most part person-parts are not systematically incorporated into perception predicates in Dalabon (cf. Ponsonnet 2015, p.122-123).

The **volition** subdomain is dominated by the verb *djare* ‘like, love, want’, (cf. Ponsonnet 2014b, p.165-169) though there are a number of related compounds involving *yolh-no* ‘pep, feelings’, as well as the not fully attested partial-synonym *ngerh-waddi* ‘feel strong affection’ (lit. ‘heart/breath+be high up’; Ponsonnet 2014b, p.434).

The **emotion** subdomain has significant internal complexity that cannot be covered in any detail (cf. Ponsonnet 2014b), though there are several significant aspects which are worth overviewing here. *Kangu-no* ‘belly’, *yolh-no* ‘feelings/pep’, and *ngerl-no* ‘heart/breath’ are noted as particularly prolific person-parts for Dalabon emotion terms (Ponsonnet 2014b, Chapter 8). These are representative of the grouping of non-visible parts and abstract attributes, which appear in compounds which “systematically represent emotions as states of the person or as parts of the person” (ibid, p.239). This is in contrast to visible person-parts like *dje-no* ‘nose’, which “describe visible responses to emotions” (ibid, p.329). *Kangu-no* ‘belly’, *kodj-no* ‘head’ and *yolh-no* ‘pep, feelings’ all appear in psycho-collocations for bodily sensations, where compounds

with *kangu-no* and *kodj-no* refer sensation associated with the literal stomach/abdomen and head respectively. *Yolh-no* can be used to refer to a general "feeling" that covers both emotion and bodily sensation. For example, in *yolh-yord(mu)* "feel better" (Lit. 'pep/feelings + body fluid be released'; Ponsonnet 2014b, p.448).

4.1.2 Morphosyntactic aspects of Dalabon psycho-collocations

This section will overview aspects of Dalabon morphosyntax which are within the scope of this study (cf. §2.4). It is worth noting again that the nature of published work on experiential construction in Dalabon gravitates toward emotion. Dalabon person-parts are largely equivalent to the "animate-part noun" class defined in Ponsonnet (2015) which in this thesis are referred to as person-parts. Members of this class are defined by being frequently incorporated, bound, "display[ing] ambiguous obligatory possession", and "allow[ing] (even favour[ing]) free possessor raising whether the noun is incorporated or not" (ibid, p.21). These properties justify calling Dalabon animate-part nouns inalienable (ibid, p.21). The remainder of this section will briefly overview these features of animate-part nouns.

All person-parts in Dalabon are bound and obligatorily possessed. This means that Dalabon person-parts "*must* occur either in verb complexes or noun phrases with some material to their right" or where neither of

these is the case, with the "filling element" *-no*, which is "homophonous with the third singular possessive enclitic" (Ponsonnet 2015, p.15, *emphasis mine*).[31] Example (10) below includes instances of *kangu-no* 'belly' incorporated into the verb complex, and later as a noun phrase with possessive marking. [32]

(10) Da-h-lng-marnu-**kangu**-yowyow-mu.
2sg>3-R-BEN-**belly**-flow:REDUP-PR

Kurdu	nga-h-dudj-miyan	biyi-ngan-kah,	dja-h-yin	**kangu**-ngu
INTJ.excl	1sg-R-return-FUT	Man-1sg-POSS-LOC	2sg-R-say/do:PR	**belly**-2sgPOSS

'You'll be pleased thanks to him [you will flow from your belly for him]. Ah. I'll go back to my husband, you're thinking for yourself [your(sic?) think from your belly]'

(Ponsonnet 2014; 25520120718b_MT_000 49 [El])

For the surveyed person-parts, there is evidence of incorporation in all subdomains, but unincorporated person-parts only appear for cognition and emotion. For the examples I have collected, unincorporated person-parts appear in experiential constructions where they can take instrumental (e.g. Ponsonnet 2009, ex.10), locative (Ponsonnet 2009, ex.9; Evans 2021, ex.53, 57), or possessive marking in addition to *-no*, analysed as morphological filler in these cases possessive marking elsewhere (cf. Ponsonnet 2015, ex.31). An example with locative case marking is

31 Note there is one example in Evans (2021, ex.57), where *men(-no)* 'consciousness' appears outside of the verb complex and without *-no* or other case marking.

32 The provided context for this example is that the speaker is "commenting on [MP's] feelings when [MP] returned to Paris, where [MP's] husband lived"

demonstrated in (11) below with *kodj-no* 'belly'. Whether cases like (11) should be understood as psycho-collocation is returned to in Chapter 5.

(11) kah-bengkan **Kodj-no-kah**
3sg/3sg-knowPR **head-3sgPOSS-LOC**
'He knows it in his head.'
Free translation: 'He's got it in mind.' (Ponsonnet 2009, p.21)

Ponsonnet defines possessor raising as when "the possessor of a noun is treated as an argument of the verb that describes an event involving the part, rather than as a simple possessor" (Ponsonnet 2014, p.70; cf. Ponsonnet 2015, p.17). Possessor raising is obligatory for incorporated person-parts, and highly frequent but non-obligatory for unincorporated person-parts (Ponsonnet 2015, p.24, 255). In Dalabon, possessor raising occurs in cross-referencing on the verb. For example, in (12) the prefix *nga-* agrees with the person of the possessor (1st person) rather than the possessum (3rd person).[33]

(12) nunda ke bibbi-ngan **nga**-h-ngarrk-mu
DEM EMPH breast-1sgPOSS **1sg**-R-ache-mu
'This is when my breast hurts' (Ponsonnet 2015, p.24; 250909_85OK 488 (LB) [ContEl])

33 The provided context for this example is that the speaker is "Discussing sensations triggered by empathy with suffering relatives, here a woman's children"

4.2 Ewe

Ewe is a Kwa (Niger-Congo) language of West Africa where “bodily expressions are the unmarked or basic expression for an emotion” in Ewe (Ameka 2002, p.29). This is the case for “Ewe and many other Ghanaian languages use them as basic terms for describing experiences without any suggestion of figurative language” (Ameka 2002, p.51-2). Similar “bodily expressions” can also be found beyond emotion, though it is less clear that they are unmarked. For example, in expressions of volition and bodily sensation (Awuku 2006) as well as for perception and cognition (Sevor 2015). I draw primarily from work reported in Awuku (2006), Ameka (1990, 1991, 1996, 2002), and Sevor (2015, Chapter 5). A key aspect of psycho-collocations in Ewe is the division between cases where the person-part appears in a phrasal predicate and those where the person-part appears in a possessive construction.

4.2.1 Person-parts in Ewe Psycho-collocations

Based on examples in Awuku (2006), Ameka (1990, 1996, 2002), and Sevor (2015, Chapter Five), it appears that the person-parts with the largest number of unique noun-verb pairings are the following (definitions follow Awuku 2006 unless specified):

Table 4.3. Ewe person-parts in database

Person-part	***Definition***
dzi	“heart”
ɖɔme	“stomach”
lãme	“flesh inside”
ŋkú	“‘eye’ as a physical organ” (Ameka 2002, p.31)
ŋu	“flesh”
ŋù	“‘eye’ as a “psychologised” organ, so to speak” (Ameka 2002, p.31)
ŋutsi	“skin”
susu	“mind, thought” (Sevor 2015)
to	“ear” (Sevor 2015)

Note that *ŋkú* and *ŋù* are related etymologically, though their use shows that they differ in their typical referent (Ameka 2002, p.31-2). This has implications for the subdomains where *ŋkú* and *ŋù* most often occur. Examples of psycho-collocation involving these person-parts distribute across the different experiential subdomains as per Table 4.4.

Table 4.4. Person-parts and subdomains in a sample of Ewe Psycho-collocations (unique noun-mate pairings)

Subdomains / *Person-Part*	***Cognition***	***Perception***	***Volition***	***Emotion***	***Bodily Sensation***
dzi "heart"				✓✓	✓✓
ɖɔme "stomach"				✓✓	✓✓
lãme "flesh inside"				✓✓	✓✓
ŋkú "eye (physical)"	✓	✓✓	✓	✓	✓
ŋu "flesh"				✓✓	✓✓
ŋù "eye (psychological)"	✓	✓	✓	✓✓	
ŋutsi "skin"					✓✓
susu "mind"	✓✓				
to "ear"	✓	✓✓		✓ *	

✓ - At least one unique mate-noun pair.
✓✓ - More than three unique mate-noun pairs.
***** Appears with *tome* 'ear.inside'

With respect to **cognition**, Sevor (2015) gives a number of examples in which thinking is associated with *susu* 'mind' such as (13). Note though that in examples given by Sevor, *susu* 'mind' always appeared as either a possessed nominal, or as an oblique argument. Several other person-parts appear in the cognition subdomain with a low number of unique mate-noun pairs: *ŋkú* "eye (physical)", *ŋù* "eye (psychological)" and *to* "ear". The Ewe word *to* 'ear' largely appears in the **perception**

subdomain, with examples for cognition and perception where it appears with the postposition *-me* ‘inside’ (cf. Ameka 1996, p.172-173).

(13) e-ʃe susu-wo yi ʊiĩ
3SG-POSS mind-PL go far
“He was lost in thought” (lit. his thoughts went far away.) (Sevor 2015, p.102)

ŋkú and *ŋù* occur across a number of subdomains. While they are interchangeable in some constructions, this is not the case in others, particularly those in the perception subdomain for *ŋkú* ‘physical eye’, as in (14). While psycho-collocations for the **volition** subdomain are largely not represented in the surveyed sources, some examples in Ameka (2002) like (15) can be interpreted as part of both the emotion and volition subdomains.

(14) me-ga-tró ŋkú ɖé-m o
2SG.NEG-REP-spin eye ALL-1SG NEG
“don’t eye me” (Ameka 1996, p.32)

(15) ʊɑ ŋù
move eye
“to be jealous or envious of someone” (Ameka 1996, p.28)

The **emotion** subdomain allows a large number of person-parts, with *ŋu* “flesh” and *dzi* “heart” occurring the most frequently for the sources surveyed. With respect to the **bodily sensation** subdomain, there are constructions largely include number of parts that act as the locus of the

sensation as in (16) where a number of person-parts can replace *lãme* "flesh inside" as the locus of pain.

(16) lãme nye vé-m
flesh.inside me pain-PROG
"I have body pains" (Awuku 2006, p.289)

4.2.2 Morphosyntactic aspects of Ewe psycho-collocations

Ewe psycho-collocations can take several different forms, I briefly overview these types of psycho-collocation here (for details see especially Ameka 1991, Chapters 7 and 10; Awuku 2006).

Phrasal predicates are one of four structural types of verbal predicate in Ewe (Ameka 1996, p.788). They are composed of "a verb root and an inherent nominal complement which together express a predicate meaning" (Ameka 1991, p.350). Ewe phrasal predicates are related to a broader class of "'inherent complement verbs', which are much disputed in the literature on Ewe, and the Kwa branch in general" (Awuku 2006, p.140-141). Ameka notes that phrasal predicates "should not … be thought of as incorporation (although it feels like it)" (Ameka 1991, p.346, 1996, p.788). Awuku (2006, p.139-140) gives restrictions on modifying the nominal component and the inability to extract the nominal using questions as evidence for the two components as "a single semantically intransitive predicate". Ewe phrasal predicates can be divided into two types based on order of the verbal and nominal components as

demonstrated by (17-18) below. These two types of phrasal predicates denote different kinds of experiential situation and are associated with different morphosyntactic properties (cf. Ameka 1990). With respect to examples in the surveyed sources, phrasal predicates only denote experiential situations in the bodily sensation and emotion subdomains.

(17) $[\text{N V}]_{\text{PRED}}$
dzi **kú**-m
heart **die**-1SG
'I was angry.' (Ameka 1990, p.148)

(18) $[\text{V N}]_{\text{PRED}}$
é-**kú** **dzi** ná-m
3SG-**die** **heart** to-1SG
'It angered me.' (Ameka 1990:156)

Note that (17-18) also demonstrate different numbers of situation members. In (17), the person-part *dzi* 'heart' is an argument of the verb *kú* 'die' and the experiencer is denoted by object marking on the verb (cf. Ameka 1990, p.148-9). This is in contrast to (18) where a stimulus subject appears in marking on the verb and the experiencer is realised as an oblique marked by the dative *ná* (cf. Ameka 1990, p.167-171). Sometimes a noun-verb pairing will only allow the arguments in specific positions while other cases, like (17-18) allow multiple configurations for a given pairing based on the situation being described (see Awuku 2006 for discussion). While these alternations are understood as possessor raising

in Ameka (1996), the appropriateness of this label for psycho-collocation is questioned by Awuku (2006, p.117-118) on the grounds that, unlike in typical possessor raising, the possessed person-part is the locus of experience. There are a handful of cases where a person-part appears in the oblique in Awuku (e.g. 2006, p.291), though these are infrequent.

Beyond phrasal predicates, person-parts also appear in possessive phrases. Ameka (1996) outlines three major kinds of nominal possession in Ewe which are associated with differing degrees of alienability. Person-parts appear in alienable possessive constructions, which are of the form POSSESSOR *fe* POSSESSED (cf. Ameka 1996, p.791-795). [34] The fact that person-parts are considered alienable in Ewe is notable, and is explained by Ameka (1996, p.810-812) with respect to functional considerations relating to homophonous pairs of words: a person-part, and a spatial relational term (e.g. *ta* 'head' and *ta* 'fore part, top, above'). In short, Ameka argues that person-parts receive a different grammatical treatment so that they can be more easily distinguished from spatial relation terms (*ibid*, p.811). Alienable nominal possession appears most frequently in the database for the cognition and perception subdomains, as demonstrated by (13) above for *susu* 'mind'. Note that while inalienable possession in Ewe is of the form POSSESSOR POSSESSED, thereby appearing similar to the phrasal predicates above,

34 The possessive *fe* is written as *ɸe* in some publications (e.g. Ameka 1996).

the person-part in these phrasal predicates cannot reasonably be interpreted as being in an inalienable possessive construction despite not appearing in inalienable possessive constructions (cf. Awuku 2006, p.138-141).

4.3 Hakha Lai[35]

Hakha Lai is a Kuki-Chin (Sino-Tibetan) language Chin State, Burma. In Hakha Lai, “psycho-collocations are essential … it is impossible to discuss the life of the mind without them” (VanBik 1998, p.227). In surveying psycho-collocations in Hakha Lai, I have primarily drawn information from Van Bik (1998, 2010), Yandt (2019) and Bickel (2000). Key aspects of psycho-collocations in Hakha Lai are (a) an apparent lack of psycho-collocations for the perception subdomain, and (b) agreement on the verb referring to the person-part rather than possessor.

4.3.1 Person-parts in Hakha Lai Psycho-collocations

The person-part nouns with the greatest number of verb pairings are defined as follows (all definitions from Van Bik 2014).[36]

35 Hakha Lai is also referred to with other names including Lai and Hakha Chin. I use Hakha Lai here primarily for consistency with Van Bik (1998, 2010) which are the most important sources on psycho-collocation in the language.

36 The spelling of person-parts in this section follows Van Bik (1998), except for examples from other sources, where the spelling from that source is used.

Table 4.5. Hakha Lai person-parts in database

Person-part	***Definition***
hmaay	“face”
hna	“ear (of a person)”
luŋ[37]	“heart”; “feeling”; “heart-liver” (Yandt 2019)
mit	“eye”
thaa	“muscle, sinew, tendon”; nerve(s) (Van Bik 1998); vein(s) (Yandt 2019)
thin	“liver”
thlaa	“spirit, ghost”
thluak	“brain”
ʔor	“throat”; “throat or pharynx” (Van Bik 1998, p.209)

Table 4.6 below shows how the above person-parts distribute across the predicate subclasses in the surveyed publications. It ought to be noted that *luŋ* ‘heart’ and *thin* ‘liver’ have far more unique psycho-collocations than the other person-parts surveyed for Hakha Lai (cf. §5.2). Van Bik (1998, p.210) notes that *luŋ* can be translated as either ‘liver’ or ‘heart’, with preferred meaning varying by region. Van Bik (*ibid.*) glosses *luŋ* as ‘heart’ on the basis of the word’s etymology, which I follow here. In some psycho-collocations, *luŋ* ‘heart’ and *thin* ‘liver’ can appear

37 Van Bik (1986, 2014) gives two headwords for *luŋ*. Whether only one is used in psycho-collocations cannot be determined from the examples given, so both definitions are included here. Yandt’s (2019) glosses suggest these could be reinterpreted as two senses under a single headword.

interchangeably (though this is not always the case), as well as being able to appear together as the compounds *luŋ-thin* and *thin-luŋ* (*ibid*).

Table 4.6. Person-parts and subdomains in a sample of Hakha Lai Psycho-collocations (unique noun-mate pairings)

Subdomains / *Person-Part*	***Cognition***	***Perception***	***Volition***	***Emotion***	***Bodily Sensation***
hmaay 'face'				✓	
hna 'ear'				✓✓	✓
luŋ 'heart'	✓✓		✓✓	✓✓	✓
mit 'eye'			✓	✓	✓
thaa 'nerves, veins'				✓	✓
thin 'liver'				✓✓	
thlaa 'spirit'				✓✓	
thluak 'brain'	✓✓				
ʔor 'throat'			✓		✓✓

✓ - At least one unique mate-noun pair.
✓✓ - More than three unique mate-noun pairs.

I will now briefly consider the person-parts used for psycho-collocations in Hakha Lai with respect to each of the subdomains in Table 4.5. The **cognition** subdomain is dominated by psycho-collocations with *luŋ* 'heart' and *thluak* 'brain' as in (19). There don't appear to be any psycho-collocations in the surveyed publications which can be understood as part of the **perception** subdomain, which is notable given *mit* 'eye' appears in other subdomains. There are three person-parts which can be

used in a psycho-collocation in the **volition** subdomain: *luŋ* ‘stomach’ *ʔor* ‘throat’ and *mit* ‘eye’, as in (20).

(19) kan thluak ʔa buay
1PL.POSS brain 3SG.SUBJ messy
‘We are confused' (lit. 'Our brain is messy') (Van Bik 1998, p.205; Yandt 2019, p.5)

(20)[38] ka mit ʔa thii
1SG.POSS eyes 3SG.OBJ die.I
I wanted/desired it/her' (lit. 'My eyes died on it/her') (Van Bik 1998, p.205);
‘I want it’ (Yandt 2019, p.14)

The **emotion** subdomain shows the greatest number and diversity of psycho-collocations, one of which is given in (21). Note that while Van Bik (1998) and Yandt (2019) note an overlap in the referents of *luŋ* ‘heart’ and *thin* ‘liver’, only *luŋ* appears in psycho-collocations outside of the emotion subdomain.

(21) ʔa thin ʔa lin
3SG.POSS liver 3SG.SUBJ hot.I
‘He is worried [helplessly and angrily]' (lit. 'His liver is hot/burning') (Van Bik 1998, p.216);
‘He feels worried, helpless, and angry’ (Yandt 2019, p.16)

With respect to **bodily sensations**, a number of kinds of bodily sensations are represented though there are a notably greater number of distinct psycho-collocations relating to hunger, especially with *ʔor* ‘throat’ as in (22).

38 Given Yandt (2019) collected alternate definitions for many of the psycho-collocations in Van Bik (1998), translations from both publications are given where they diverge. Glossing follows Van Bik (1998) unless otherwise stated.

(22) ʔa ʔor ʔa vuay
3SG.POSS throat 3SG.SUBJ satiate.I
‘He/She is satiated' (lit. 'His/Her throat is satisfied') (Van Bik 1998, p.210);
‘He ate enough’ (Yandt 2019, p.8)

4.3.2 Morphosyntactic aspects of Hakha Lai psycho-collocations

This section briefly overviews several aspects of the morphosyntax of psycho-collocations in Hakha Lai. In particular, it focusses on the significance of person and number marking on verbs, and possessive marking on person-parts. The reader is referred to Bickel (2000) and Van Bik (2010) for greater detail than is given here.

Verbs in Hakha Lai demonstrate modified stems (apophany), which are for the most part associated with different transitivity (Peterson, 2014, p.262). Almost all psycho-collocations involving person-parts employ what linguists have labelled form I stems, which correspond to less transitive actions (Peterson, 1998, p.89). While transitive constructions do occur for psycho-collocations, they require “explicit, derivational transitivisation” rather than different forms of the verb stem (Bickel 2000, p.589).

Although form I verbs typically take a single participant, this is not the case with form I verbs in psycho-collocations, which allow interpretations with either one or two participants (Van Bik 2010, p.144).

This is shown in (22), where the clitic *ʔa-* can used for either 3SG.S and 3SG.O. Both transitive and intransitive interpretations are allowed in (22), but this is not possible in constructions with form I verbs that are not psycho-collocations as demonstrated by (24a-b) (Van Bik 2010, p.144). Note that the clitic *ʔa-* does not agree with the possessor in either interpretation of (23), which is in the 1st person.

(23) ka-lûŋ$_{i}$ **ʔa$_{i/j}$-r̥ook**
1SG.POSS-heart **3SG.S/3SG.O**-break.down.I
a. 'I am disappointed'
b. 'He disappoints me'
(Van Bik 2010, p.143-4)

(24) a. ka-tsòo-lèeŋ ʔa-rook
1SG.POSS-cow-cart 3SG.S-break.down.I
'My bullock cart broke down' (Van Bik 2010, p.143)

b. *ka-tsòo-lèeŋ na-rook
1SG.POSS-cow-cart 2SG.S-break.down.I
*'You broke my bullock cart' (Van Bik 2010, p.144)

Van Bik (2010) argues that these cases include the 3SG.O null-morpheme which agrees with the experiencer. This is shown in (25). Given 1st person clitics in the same position are not null-morphs in Hakha Lai (cf. Van Bik 2010, p.139), Van Bik's view might be more accurately interpreted as the null-morph verbal clitic agreeing with the person-part, which is metonymically associated with the experiencer.

(25) kan-lûŋ$_i$ nan$_j$-**Ø$_i$**-rook
3SG.POSS-heart 2SG.S-**3SG.O**-break down.I
'You (pl) disappoint us.' (Van Bik 2010, p.145)

Bickel notes that this is different from what is typically meant by agreement, whereby the "referential features marked by nominals and verbal agreement markers", in this case person and number marking, demonstrate the status of the nominal as an argument of the verb ("*identificational* agreement"; Bickel 2000, p.584). This is distinct from Hakha Lai psycho-collocations, whereby agreement "indexes the person **with regard to whom** the psycho-collocation holds" ("*relational* agreement"; Bickel 2000, p.588, bold in original).

Van Bik (2010, p.146-8) demonstrates this again with respect to form III verbs, which always have a causative meaning. In this context, changing the possessor of the person-part then changes who is understood as the experiencer, without any concomitant change in marking on the verb. For example, in (26) *na-* marks the 2nd person singular participant (later marked as O on the verb) as the experiencer. This is in contrast to (27) below, where *ka-* marks the 1st person singular participant as the experiencer (later marked as S on the verb). This kind of syntactic behaviour is only found in psycho-collocations (Van Bik 2010, p.147-8). Note that a similar argument is given for form II verbs in Bickel (2000, p.588-591).

(26) nàŋmaʔ **na**-lûŋ ka-**ń**-hrinʔ
2SG.PRON **2SG.POSS**-heart 1SG.S-**2SG.O**-make.green.III
‘I make you suspect me’ (Van Bik 2010, p.146)

(27) nàŋmaʔ **ka**-lûŋ **ka**-ń-hrinʔ
2SG.PRON **1SG.POSS**-heart **1SG.S**-2SG.O-make green.III
‘I suspect you’ (Van Bik 2010, p.147)

Van Bik (2021, p.381) notes that psycho-collocations are part of a larger class of “N-V collocations” which occur in many Kuki-Chin languages, in which “some verbs occur only with a particular noun”. Presumably this means that some verbs can only allow certain interpretations when they occur with particular nouns, given the examples provided on the same page all include verbs which occur without those nouns, and with different interpretations (following examples given in Van Bik 1986). It is unclear whether all N-V collocations share the above properties for psycho-collocations discussed in Bickel (2000) and Van Bik (2010).

4.4 Mangap-Mbula[39]

Psycho-collocation appears in Mangap-Mbula, an Oceanic (Austronesian) language of the Morobe Province of Papua New Guinea. Bugenhagen notes that "body image expressions … are certainly the most productive and semantically richest means of talking about experiences [in Mangap-Mbula]" (2001, p.74) to the extent that speakers "must use body image expressions which localise the experience" (*ibid*, p.75). Psycho-collocations are the dominant means for coding emotional and physical experiences, given there are relatively few cases in which an individual lexeme can encode an experiential situation in these subdomains (ibid). This section primarily draws from information in Bugenhagen (1990, 1995, 2001, 2002) and Bugenhagen & Bugenhagen (2007).

4.4.1 Person-parts in Mangap-Mbula Psycho-collocations

There are many person-parts which are used in Mangap-Mbula psycho-collocations, the most productive of these are defined by Bugenhagen (1990, 1995). The below definitions are ordered with respect to productivity in psycho-collocations as given in Bugenhagen (1990:204; definitions refer to Bugenhagen 2001 unless specified): [40]

39 Note that this language has also been referred to with other names, such as Maŋaaba-Mbula (cf. Bugenhagen 1995, p.13), and that Bugenhagen uses the term "body image expression" (in a way distinct from the use of the term in Wierzbicka 1999) for what I refer to here as psycho-collocations.

40 Note that there are several other person-parts which appear in psycho-collocations as described by Bugenhagen (1991; 1995; 2001). The person-parts also appeared a single time in my survey, and are not considered.

Table 4.7. Mangap-Mbula person-parts in database

Person-part	***Definition***
mata-	"Various senses include 'eye, face, front, presence, edge of a sharp object, first, chief, brightness'" (p.77).
lele-	"part or space located inside of something, insides/feelings" (p.86).
kete-	"liver, chest, flat-shaped part of something" (p.95).
ni-	"somewhat difficult to define. It is used to talk about the physical sensations which are not localised in any one particular part of the body... *Ni-* is not, however, equivalent to *body* in English" (p.99)
kuli-	"skin, bark, surface" (p.101)
kopo-	"stomach, belly, cavity" (p.104)
talŋa-	"ear" (Bugenhagen and Bugenhagen 2007:504)
kuzu-	"nose" (Bugenhagen and Bugenhagen 2007:155)

The high productivity of *lele- and kete-* in Mangap-Mbula mirrors a more general Oceanic pattern where two analogous lexemes (reconstructed for proto-Oceanic as **qate-* and **lalom*) are used in psycho-collocations (Fox, 2018, p.10-19; Osmond, 2016). Several of the definitions given in Bugenhagen (2001) are decanted into multiple headwords in a later dictionary by Bugenhagen and Bugenhagen (2007), which are presented below. Entries for *mata-*, and *ni-* include two headwords which are involved in psycho-collocation: one which refers to a concrete person-part, and another which refers to an incorporeal person-part. *kete-* has three relevant headwords (note only one sense of the *kete-*$_1$ headword appears to be relevant, which is listed):

mata-₁	"eye (physical organ of sight)" (2007, p.196)	***ni-₄***	"body, part of the body that experiences physical sensations" (2007, p.290)[41]	***kete-₁***	"chest, internal organs (general area, can refer to lungs, stomach, etc.)" (2007, p.108)
mata-₄	"mind (the part of a person that thinks and wills)" (2007, p. 200)	***ni-₆***	"emotional centre for some feelings" (2007, p.290)	***kete-₂***	"heart (organ)" (2007, p.109)
				kete-₃	"seat of strong emotions (mainly used for various types of uncontrolled feeling)" (2007, p.109)

These person-parts distribute across experiential subdomains as summarised in Table 4.8.. Note that it is not feasible to differentiate the various senses of each person-part given in Bugenhagen and Bugenhagen (2007).

41 Note that there are important differences between *ni-₄* and English *body*, for example regarding dead bodies (Bugenhagen 2001, p.99).

Table 4.8. Person-parts and subdomains in a sample of Mangap-Mbula Psycho-collocations (unique noun-mate pairings)

Subdomains / *Person-Part*	***Cognition***	***Perception***	***Volition***	***Emotion***	***Bodily Sensation***
***mata-*‘eye, mind’**	✓✓	✓✓	✓✓	✓✓	✓✓
***lele-* ‘insides, feelings’**	✓✓*		✓✓	✓✓	✓
***kete-* ‘heart, internal organs’**	✓**		✓✓	✓✓	✓✓
***ni-* ‘body, emotional centre’**	✓***		✓✓	✓✓	✓✓
***kuli-* ‘skin’**	✓***	✓	✓	✓✓	✓✓
***kopo-* ‘stomach’**				✓✓	✓✓
***talŋa-* ‘ear’**	✓	✓✓	✓		✓✓

✓ - At least one example.
✓✓ - More than three examples.
***** - Doubt
****** - Absence of thinking
******* - Knowing another person intimately

I will now briefly consider the person-parts used for psycho-collocations with respect to each of the subdomains of the experiential situation given in Table 4.8. First considering the **cognition** subdomain, Bugenhagen (2001, p.85) notes that “no other person-part is so frequently associated with expressions indicating thinking and knowing” as *mata-* ‘eye’, though *lele-* ‘insides’ is associated with doubt and both *ni-* ‘emotional centre for some feelings’ and *kuli-* ‘skin, bark,

surface' are associated with intimate knowledge of another person.[42] The **perception** subdomain is dominated by person-parts referring to sensory organs which are more directly involved in perception (i.e. for *mata-* 'eye', *kuli-* 'skin', and *talŋa-* 'ear'). **Volition** is dominated by *lele-*, though it should be noted that *mata-*$_{4}$ and *ni-*$_{6}$ can be used when referring to something that is not intended or desired. There are many facets of what I have labelled **emotion**, and as such there are many person-parts used for emotion predicates; *lele-* 'insides', *kete-*$_{3}$ 'seat of strong emotions' and *ni-*$_{6}$ 'emotional centre' appear to be particularly productive. Distinctions between Mangap-Mbula emotions are explored in more detail in Bugenhagen (2001). **Bodily sensations** are spread across a number of person-parts where they are used to refer to a perceived location of sensation as in (28). Note that while Mangap-Mbula has a very small set of experiential verbs, these verbs broadly cover all of the above subdomains (Bugenhagen 1995, p.255).

(28) **ni-**ŋ i-tek-teege
body-1S.PSR 3S-hold-RED
"I am shivering with cold" (Bugenhagen 1990, p.201; cf. Bugenhagen 2002, p.15)

42 Referring to a situation where one person knows another well enough that they "therefore [do] not feel embarrassed or afraid around them" (cf. Bugenhagen 2001, p.104; Bugenhagen and Bugenhagen 2007, p.291)

4.4.2 Morphosyntactic aspects of Mangap-Mbula psycho-collocations

This section will overview aspects of Mangap-Mbula morphosyntax which are within the scope of this study, following the list of aspects of psycho-collocations at the end of Section 2.4.

Mangap-Mbula distinguishes alienable and inalienable nouns, where inalienable nouns take genitive case marking (Bugenhagen 1995, p.90-98, 202-203).[43] All of the above person-parts are listed as inalienable nouns in Bugenhagen and Bugenhagen (2007). The genitive case is marked with a postnominal suffix which inflects for number, person and clusivity in agreement with a non-obligatorily expressed experiencer (Bugenhagen 1995, p.91). This is consistent with direct possessives found in many Oceanic languages, as described in (Lichtenberk 1985, p.93). The genitive suffix on inalienable nouns is analysed as formally marking the existence of a part-whole relationship (Bugenhagen 1995, p.93, 379). Example (29) below demonstrates this.

(29)	n-io	kete-**ŋ**	malmal	pi-no
	GIV-1S	Liver-**1s.PSR**	fight	REF-3S

'I am angry at him.' (Bugenhagen 1990, p.202)

Person-parts can form a complex predicate with the verb with the following structure (Bugenhagen 1995, p.256):

43 Mangap-Mbula makes three additional divisions within alienable genitives, but this does not appear to be relevant for psycho-collocations.

[NP]$_{NP\text{-}SUBJECT}$ [PERSON.PART-GEN VERB X]$_{PREDP}$

Bugenhagen (1990, 1995) argues for this on the basis of allowed positions of modal auxiliaries (Bugenhagen 1995, p.257). For example, *ko*, an auxiliary indicating uncertainty, cannot be inserted between the person-part and the verb in (30). Note that this condition does not hold for person-parts which are used in non-experiential contexts as seen in (31).

(30) (ko) zin (ko) lele-n **(**ko)** i-saana
UC NOM.3PL UC insides-GEN.3PL UC 3SG-deteriorate
'They might/will be upset' (Bugenhagen 1995, p.257)

(31) (ko) ruumu **(**ko)** ute-ene (ko) i-saana
UC house UC head.GEN.3SG UC 3SG-deteriorate
'The roof of the house might/will be bad' (Bugenhagen 1995, p.257)

Notable edge-cases among Mangap-Mbula psycho-collocations come from the volitional predicates including *lele-*, as seen in examples (32-33) below. Note that there is no (overt) verbal component, and that *lele-* appears to be analysed as the predicate in some publications (e.g. Bugenhagen 2002 in (33) below). Given verbs in Mangap-Mbula psycho-collocations either have a 3rd person singular prefix, or no prefix at all (1995, p.256, 257),[44] it is difficult to determine whether *lele-* is being treated as a verb in these cases. While Bugenhagen (1995, p.255) notes that "[the] inalienable body-part noun [...] may or may not be followed by a concrete verb", this optionality seems to be limited to volitional

44 Note that for the most part Mangap-Mbula verbs have a prefix that agrees in person and number with the subject (Bugenhagen 1995, p.116-117)

predicates including *lele-* in my sample. Additionally, it also appears these constructions require the experiencer to be explicit among the sampled predicates in the volition class.

(32) nio **lele-**ŋ pini
NOM.1SG **insides-**GEN.1SG REF.3SG
'I like him' (Bugenhagen 1995, p.386)

(33) nio **lele-**ŋ=pa koroŋ=ta
1SG.NOM **want-**1SG.GEN Something[??]
'I want something' (Bugenhagen 2002, p.13)

4.5 Yucatec Maya

Yucatec Maya is a Yucatecan language of the Yucatán Peninsula in Mexico and nearby regions in Guatemala and Belize. The language is referred to as Yucatec Maya by linguists and as Maya (as well as *maya?-t'aan* 'Maya language') by speakers (Bricker et al 1998:ix). Yucatec Maya "uses person and body parts in figurative expressions in rendering experience, especially emotional experience" (Verhoeven 2007:340). I primarily follow Verhoeven's (2007, Chapter 5, 6, and 8) comprehensive description of experiential constructions in Yucatec Maya, with reference to the Maya-English dictionary by Bricker et. al (1998). A key aspect of psycho-collocation in Yucatec Maya is the dominance of *óol*, which functions as "a kind of desemanticized experiential marker." (Verhoeven 2007, p.10).

4.5.1 Person-parts in Yucatec Maya Psycho-collocations

Verhoeven (2007, Chapter 8) identifies eight person-parts as significant for psycho-collocation in Yucatec Maya. These are defined as follows in Bricker et al (1998).[45] Note *pòol* and *ho'l* are treated as a single person-part following Verheoven (2007).

45 Spelling follows the Bielefeld/Erfurt Yucatec Maya orthography as applied by Verhoeven (2007, p.7-8).

Table 4.9. Yucatec Maya person-parts in database

Person-part	***Definition***
chi'	"mouth, edge" (p.68)
ho'l	"head, hair" (p.110)
ich	"face, eye" (p.10)
ìik'	"wind, air, breath" (p.12)
ni'	"nose, tip, point" (p.196)
óol	"heart, will, energy, spirit" (p.17); "pleasure, strength, energy [when referring to bodily sensations]". (p.330)
pòol	"head" (p.219)
puksi'k'al	"heart" (p.223); "… the material or immaterial body part 'heart' or to the stomach as a material body part [in the dialect investigated by Verhoeven]" (Verhoeven 2007, p.169)
xikin	"ear" (p.258)

Given *óol* dominates psycho-collocations in Yucatec Maya, it is worth further characterising the word. Speakers are only able to give a meaning of *óol* that is distinct from psycho-collocation with difficulty, with Verhoeven (2007, p.330-333) interpreting the term as having different meanings depending on the experiential subdomain. In contrast, when discussing "adjectival experiential constructions", (Lehmann 2015, p.1432) notes that "even mosquitoes have an *óol*, so there is little room for psychological speculation on the Yucatec construction". Nonetheless, *óol* is included here given it is understood as a person-part in other analyses (e.g. Lehmann 2002) and is part of a significant number of experiential constructions in Yucatec Maya.

Table 4.5 shows how the above person-parts distribute across the predicate subclasses for psycho-collocations in Verhoeven (2007). Table 4.10. agrees with claims about distribution across subdomains in Verhoeven (2007, Chapter 8).

Table 4.10. Person-parts and subdomains in a sample of Yucatec Maya Psycho-collocations (unique noun-mate pairings)

Subdomains / *Person-Part*	***Cognition***	***Perception***	***Volition***	***Emotion***	***Bodily Sensation***
chi' 'mouth'		✓✓	✓*	✓	
ìik' 'air'	✓✓				✓
ich 'eye'		✓	✓*	(✓**)	
ni' 'nose'		✓			
óol 'mind'	✓✓		✓✓	✓✓	✓✓
pòol/ho'l 'head'	✓✓				
puksi'k'al 'heart, stomach'				✓✓	✓✓
xikin 'ear'	✓	✓✓			✓

✓ - At least one example.
✓✓ - More than three examples.
***** Only volitional in addition to perception
****** Only in derivation of person-part which functions as predicate

I will now briefly consider the person-parts used in Yucatec Maya psycho-collocations with respect to each of the subdomains in Table 4.5. The **cognition** subdomain is "characterized by person-part collocations that can be distinguished from emotional ones by being less related to

óol [‘mind’]” (Verhoeven 2007, p.265). While *óol* is still present, the subdomain also includes person-parts like *pòol* ‘head’ in (34).

(34) hach h p’áat t-in **pòol**
really PFV stay(CMPL) LOC- POSS.1.SG **head**
‘I kept it in my mind’ (Verhoeven 2007, p.219; CPP_0021)

Psycho-collocations in the **perception** subdomain are characterised by the use of sensory organs as the person-part as in (35). These person-parts include *chi’* ‘mouth’, *ich* ‘eye’, *ni’* ‘nose’ and *xikin* ‘ear’.

(35) Nip’**chi’**-t a w-u’y-eh!
taste:**mouth**-TRR(IMP) SBJ.2 0-feel-SUBJ
‘Try it so that you taste it!’ (Verhoeven 2007, p.236; RMC_1673)

The **volition** subdomain largely includes psycho-collocations with *óol* ‘mind’, as in (36).[46] Additionally, there are a few examples with *chi’* ‘mouth’ and *ich* ‘eye’ relating to both enjoying perception and liking in general (cf. Verhoeven 2007, p.180).[47]

(36) ts’íib**óol**-t
write:**soul**-TRR
‘wish, desire’ (Verhoeven 2007, p.222)

46 While *óol* is glossed as ‘soul’ in (36), surrounding context suggests this refers to the same word as *óol* ‘mind’.

47 The interpretation I give (that these examples also refer to liking in general) may not appear to be entirely clear. Although they include “liking” as part of their definition (separated by semicolons for those with *chi’* ‘mouth’ and a comma for *ich* ‘eye’) in Verhoeven (2007, p.180), they are classified as perception, rather than both perception and volition. I’ve ended up deciding to include these in the volition subdomain on the basis that I included similar examples in the volition subdomain when discussing other languages and doing so does not seem contradictory to the subdomains given in Verhoeven (2007).

The **emotion** subdomain is also "characterised by person-part collocations" (Verhoeven 2007, p.263). In contrast to cognition, the emotion subdomain is dominated by psycho-collocations with *óol* 'mind' like (37) (ibid). Relative to *óol* 'mind', *puksi'k'al* 'heart, stomach' occurs marginally in the emotion subdomain, as well as possibly being a calque of Spanish expressions involving the heart in this context (Verhoeven 2007, p.263).

(37)	tèen-e'	hach	háak'-a'n	in	w-**óol**
	me-TOP	really	scare-RSLTV	POSS.1.SG	0-**mind**

'I am glad / scared' (Verhoeven 2007, p.166)

The **bodily sensation** subdomain includes psycho-collocations with both physical person-parts and *óol* 'mind' used to refer to the person as a whole (Verhoeven 2007, p.263). For the sampled person-parts, *puksi'k'al* 'heart, stomach' is the only example of a psycho-collocation with a corporeal person-part, as in (38).

(38)	hach	k-u	y-èel-el	in	**puksi'k'al**
	really	IMPF-SBJ.3	0-burn-INCMPL	POSS.1.SG	**heart**

'my stomach is burning' (Verhoeven 2007, p.336; NMP_0353)

4.5.2 Morphosyntactic aspects of Yucatec Maya psycho-collocations

This subsection briefly notes some aspects of the morphosyntax of psycho-collocations in Yucatec Maya. Given there is more depth to this topic than can be covered here, I focus on noting types of psycho-collocation and refer the reader to Verhoeven (2007) for greater detail.

Yucatec Maya psycho-collocation can first be divided into those where the person-part is either compounded or incorporated with the predicate (labelled phrasal compounds) and those where the person-part is independent of but required by the predicate (labelled collocations). Note that there is no "regular syntactic operation" between these two types (Verhoeven 2007, p.280). While there are also some cases where the person-part acts as the predicate itself like *óot* 'agree; wish, want' mentioned above, these will be put aside here.

Examples of a psycho-collocations with an incorporated person-part are given in (35-36) above, and a compounded person-part with a predicative adjective in (39) below. In all but a handful of cases, these compounds represent lexicalisations of collocations in which the person-part is not part of the verb complex (Verhoeven 2007, p.158).

(39) **síis-óol** in wíinklil
icy-mind POSS.1.SG body
'I feel fresh.' (Verhoeven 2007, p.172; RMC)

Note this classification does require that the person-part is argument of the verb, appearing for example in (34) above, where the person-part appears in a prepositional phrase with locative marking on the possessor clitic *in* 'my'. Verhoeven (2007) divides experiential collocations into those in which the person-part is part of a complex predicate and those in which it is not, i.e. whether the experiential collocation is idiomaticized. This distinction can be demonstrated by

considering whether the person-part can be topicalized. For example, *óol* ‘mind’ in (40a) cannot be separated from the predicative adjective using topicalization. Instead, the experiencer is topicalised by adding an additional pronoun as in (40b). Considering this, *óol* ‘mind’ is understood as part of the larger predicate (Verhoeven 2007, p.166). Note that the majority of idiomatized experiential collocations involve *óol* ‘mind’ (Verhoeven 2007, p.356).

(40) a. *****in w-óol-**e’** hach ki’mak/ háak’-a’n etc.
POSS.1.SG 0-mind-**TOP** really happy/ scare-RSLTV
intended: ‘as for my heart, it is really happy/scared etc.’ (Verhoeven 2007, p.166; ACC)

b. **tèen-e’** hach ki’mak/ háak’- a’n in w-óol
me-TOP really happy/ scare-RSLTV POSS.1.SG 0-mind
‘I am glad / scared’ (Verhoeven 2007, p.166)

4.6 Overview

This chapter has overviewed the five languages which are the focus of this study (Dalabon, Ewe, Hakha Lai, Mangap-Mbula and Yucatec Maya; cf. §3.2) with respect to the database introduced in Section 3.3. For each language I noted the person-parts which appear in the database and overviewed how those person-parts distribute over the subdomains of the experiential situation given by Verhoeven (2007; cf. §2.1.3). In addition, I overviewed aspects of each language's morphosyntax which can be understood as differentiating some or all of the database entries collected for a language from other constructions.

Returning to the application of the notion of psycho-collocation beyond MSEA, the overviews given in Chapter 4 suggest the potential that these languages demonstrate broad psycho-collocation as introduced in Section 2.2.1 and defined in 2.3.2. Note that this may not be a very high bar to clear, given Wierzbicka's analysis of emotion language implies that this definition is satisfied by *all* languages with respect to emotion (Wierzbicka, 1999, p.276). Similarly, Matisoff (1986, p.3-4) and later authors like Enfield (2021, p.222) note that languages across the world demonstrate something like psycho-collocation, but that mainland Southeast Asian languages do so in a way that is qualitatively different. While Section 2.2.3 introduced the idea that authors beyond MSEA have noticed and investigated phenomena which is analogous to psycho-

collocation, Chapter 4 goes beyond this. It does so by more systematically laying out how, for each language under discussion, authors have noted that something analogous to broad psycho-collocation is a *productive* means of talking about experiential constructions for some kinds of experiential situation, particularly emotion. In the chapters that follow I investigate the potential that some or all of these constructions can be understood as examples of psycho-collocation by way of comparison between the languages considered in this thesis.

5 General analysis toward a refined definition of psycho-collocation

This chapter aims to explore what constitutes typical and non-typical psycho-collocation with respect to three areas which are of relevance to the initial definition of psycho-collocation given in Section 2.3.2. Section 5.1 considers the subdomains of the experiential construction introduced in Section 2.1.3, Section 5.2 looks more closely at a possible grouping of person-parts with respect to those subdomains. Section 5.3 applies Bickel's (1997) typology of the "possessive of experience" to investigate whether psycho-collocations are lexemic phrasemes as introduced in Section 2.3. Each of these subsections draws on the database introduced in Chapter 3. In doing so, this chapter goes beyond the consideration of individual languages in Chapter 4, and looks towards comparison between those languages. This leads to a revised preliminary definition of psycho-collocation, which is given in Section 5.4. In constructing this definition, this chapter aims to address the second research question given in Chapter 1: to deepen the notion of psycho-collocation so that it can be used in comparative study. This lays the groundwork for Chapter 6, which considers the notion of psycho-collocation with respect to other kinds of terms used in comparative linguistics, as well as noting the limitations of this thesis and areas for future research.

5.1 Psycho-collocations as experiential constructions

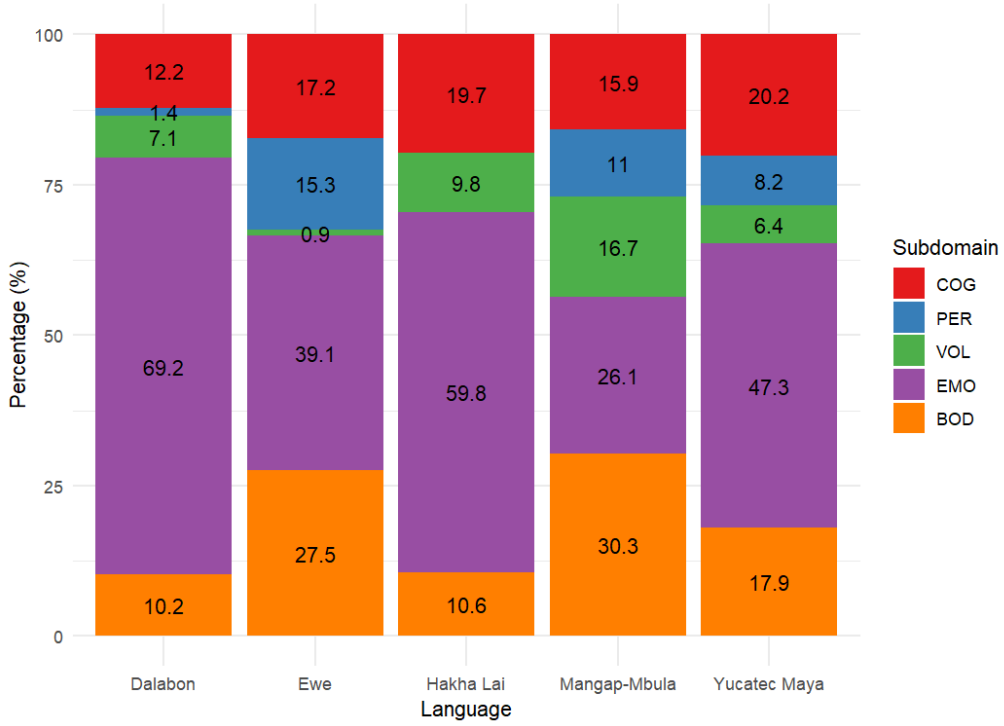
In this section I consider the application of the experiential subdomains given by Verhoeven (2007; cf. §2.1.3) to entries in the database introduced in Chapter 3. I do so by first considering the distributions of each subdomain in Section 5.1.1. Then, in Section 5.1.2 I turn to a small set of experiential meanings used in the Natural Semantic Metalanguage and the Valency Patterns Leipzig Online Database to question whether the distribution of database entries discussed in Section 5.1.1 is indicative of terms used to express experiential meanings by languages in the database. I then consider three kinds of meanings which appear in databases that go beyond the experiential subdomains in Section 5.1.2, ultimately questioning the role of those subdomains in a revised definition of psycho-collocation.

5.1.1 Experiential subdomains in the database

I begin with a general overview of the experiential subdomains introduced in Section 2.1.3 across all of the languages in the database (cf. §3.2). Each entry in the database was coded for one or more of the subdomains. I give the proportion of unique pairings between psycho-noun and psycho-mate for each subdomain in Figure 5.1. In generating Figure 5.1, I weighted the contribution of noun-mate pairs which were

assigned more than one subdomain given one entry could be coded for multiple subdomains.[48] Figure 5.1. then shows the percentage total database entries by subdomain, accounting for those cases where one entry was coded as more than one subdomain.

Figure 5.1. Proportion of each subdomain in the database for unique mate-noun pairings



The most significant aspect of Figure 5.1 is that, for the most part, (broad) psycho-collocation occurs in all the experiential subdomains named by Verhoeven (2007). The exception to this is Hakha Lai, where no entries in the database could reasonably be interpreted as part of the perception subdomain. This is particularly interesting considering the

48 For example, if an entry was assigned to both bodily sensation and emotion, it would be counted as 0.5 of an entry for the purpose of counting the entries in each subdomain.

person-parts *hna* ‘ear’ and *mit* ’eye’ appear in the database for Hakha Lai (e.g. in (41) below), given other “organs of perception” (corresponding to English *eye*, *ear* etc.) appear in the perception subdomain for database entries in each of the other languages (cf. §5.2 below on person-parts).

(41) *Hakha Lai*

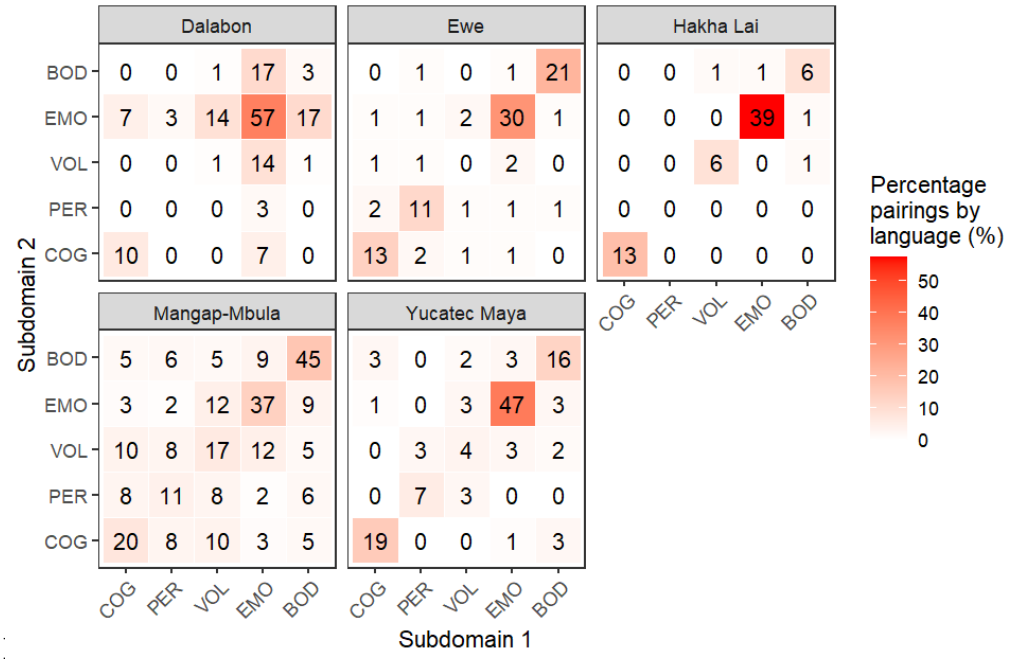
ka	**hna**	na	**hnok**
1SG.POSS	**ear**	2SG.OBJ	**entangle.I**

“You bother me” (VanBik 1998, p.205, 231; cf. Yandt 2019, p.8)

Another notable aspect of Figure 5.1 is that the emotion subdomain accounts for the greatest proportion of database entries for all languages except Mangap-Mbula, where the greatest proportion of mate-noun pairs were coded for the bodily sensation subdomain (followed by the emotion subdomain). In the other four languages under investigation, the second greatest proportion is either cognition (in Dalabon, Hakha Lai and Yucatec Maya) or bodily sensation (in Ewe). It is worth noting that the relative proportion of subdomains in Figure 5.1 might reflect a significant overlap in subdomains rather than a high proportion of that subdomain. Considering this, it is worth considering the overlap between subdomains for entries in the database. These are given Figure 5.2, which shows the number of co-occurrences for pairs of subdomains in the database.[49]

49 Note this Figure counts database entries which are coded with more than two subdomains multiple times. For example, an entry coded as perception, volition and emotion would be included in the cells for perception:volition, volition:emotion and emotion:perception. Additionally, note that the diagonal

Figure 5.2. Number of co-occurrences between subdomains for each language



subdomain. Looking across languages (cf. Table 5.1 below), the proportion of unique mate-noun pairs which were coded as a single subdomain range from 64% for Dalabon to 97% for Hakha Lai. Figure 5.2 also reiterates what was noted above: that the bodily sensation, perception, and cognition subdomains were identified most frequently across languages in the database.

Figure 5.2. also adds complexity to the picture of the subdomains given by Figure 5.1. Noting that looking at the diagonal of values for each language which indicate the database entries that were coded as a single

counts examples which are coded as a single subdomain. This means the cell for an example where "Subdomain 1" and "Subdomain 2" are both listed as cognition includes the count of unique mate-noun pairings which were *only* coded as in the cognition subdomain in a given language.

subdomain, Figure 5.2 shows two cases where languages have no examples which are coded as a single subdomain. In Dalabon, there were no examples which were coded as the only part of the perception subdomain, and in Ewe there was no example which was coded as exclusively the volition subdomain. These aspects of the database might be interpreted in a few different ways. While it might be reasoned that these results demonstrate (broad) psycho-collocation is more typical of some experiential subdomains over others, it could just as well be reasoned that the surveyed research tends toward focussing on certain experiential subdomains over others, or that some experiential subdomains tend to be associated with fewer distinct experiential constructions overall.

Table 5.1. Proportion of unique mate-noun pairs coded as a single subdomain

Language	***Proportion of database entries (%)***
Dalabon	64.0
Ewe	93.8
Hakha Lai	97.0
Mangap-Mbula	72.6
Yucatec Maya	89.4

Considering the possibility that surveyed research tends toward focussing on certain experiential subdomains over others, Section 3.3 noted that it is not reasonable to assume not all languages considered here are equally covered by available literature. This is certainly the

case with respect to Dalabon, as demonstrated above by the number of unique mate-noun pairs associated with the emotion subdomain (n=97/111, ~87%). However, Yucatec Maya and Mangap-Mbula are languages for which there is reason to believe the surveyed sources are more representative of experiential constructions more broadly, across all experiential subdomains in the database. In Yucatec Maya and Mangap-Mbula, unique mate-noun pairs are unevenly distributed across subdomains. Though, as noted above with respect to Figure 5.1, it is not the case that database entries in these two languages are unevenly distributed across subdomains in precisely the same way. Following from this, assuming representative samples of broad psycho-collocation in Yucatec Maya and Mangap-Mbula, the database suggests a tendency toward an uneven distribution of psycho-collocations across experiential subdomains which cannot be explained by the fact that existing research focusses on some subdomains over others.

Two explanations for this tendency are either that psycho-collocation itself is associated with some subdomains over others, or that some experiential subdomains have a greater diversity of experiential constructions overall. Considering other research has suggested that across languages a wider range of constructions are used for emotions than other experiential subdomains, the more parsimonious interpretation is the latter (cf. e.g. Aikhenvald et al., 2001; Aldai &

Wichmann, 2018; Luraghi, 2021, p.23; Verhoeven, 2007). Considering this, the notion that psycho-collocation is primarily associated with the emotion subdomain that appears in the definition makes up a part of the revised definition of psycho-collocation in Section 5.4.

5.1.2 Comparison of "core" meanings of experiential subdomains

Even for those examples which were only coded as a single subdomain, there is a remaining question of comparability between experiential meanings: are the distributions across experiential subdomains discussed thus far indicative of the "core" of that subdomain in the language?[50] In each of the subsections in Chapter 4 I suggested this is the case for some or all of the subdomains in passing using illustrative quotes in each introductory paragraph. However, it is worth returning to the topic here when looking across the languages in the database.

Comparative studies of experiential verbs often focus on restricted lists of comparable predicates (cf. e.g. the list given by Bossong (1998) for languages of Europe). For considering the experiential subdomains in more detail, I investigated examples relating to the predicate meanings given in Table *5.2*, which have been sorted according to experiential subdomain. It is reasoned that this is a suitable set of experiential

50 I return to a broader discussion of the question of comparability in Section 6.2.

predicates given their inclusion in either the Valency Patterns Leipzig Online Database (hence ValPaL; Hartmann et al., 2013) or as a semantic prime in the Natural Semantic Metalanguage (hence NSM; cf. Goddard & Wierzbicka, 2014). Both the ValPaL corpus and researchers who apply NSM are interested in cross-linguistically comparable meanings, though they differ in the set of meanings they investigate and the research focus following from those meanings.

Table 5.2. Select predicate meanings from NSM and ValPaL sorted by experiential subdomain[51]

		Predicate meanings	
Experiential subdomain	***Cognition***	THINK*	KNOW*#
	Perception	SEE*#	HEAR*#
	Volition	WANT*	LIKE#
	Emotion	BE SAD (ABOUT)#	FEAR#
	Bodily Sensation	BE HUNGRY#	FEEL PAIN#

***** Has a corresponding semantic prime in the Natural Semantic Metalanguage
# Appears in The Valency Patterns Leipzig Online Database (Hartman et al., 2013)

For each of the languages considered here I returned to dictionaries for each language, this time considering entries which referred to the English terms for the meanings given in Table 5.2.[52] An overview of the

51 There are several relevant meanings in ValPaL that I have not included here: SMELL, TOUCH, FEEL SICK and FEEL COLD. TOUCH also appears as an NSM prime. Additionally, note that I group FEAR with FRIGHTED and SEE with LOOK (AT) here, though these are differentiated in ValPaL.

52 These are the same dictionaries given in Table 3.2, except for Hakha Lai, where I referred to D. Van Bik (1996).

results are given in Table 5.3, which represents whether a distinct psycho-noun and psycho-mate could be identified for at least one translation.

Table 5.3. Potential psycho-collocation identified select experiential meanings

Subdomain	***Cognition***		***Perception***		***Volition***		***Emotion***		***Bodily sensation***	
Language	*THINK*	*KNOW*	*SEE*	*HEAR*	*WANT*	*LIKE*	*BE SAD*	*FEAR*	*BE HUNGRY*	*FEEL PAIN*
Dalabon	X	X[a]		[b]		X[c]	X[d]	X[e]	X	**
Ewe	X			X			X		X	**
Hakha Lai							X	X	[f]	[g]
Mangap-Mbula	X		X	X	*	*	X	X	X	**
Yucatec Maya		X			*	*	X	X		**

X Relevant dictionary contains an example where a term for one of the predicate meanings in Table 5.2 contains a distinguishable psycho-noun and psycho-mate.

***** Only examples where person-part is predicate (cf. discussion below)

****** Examples of constructions where a person-part localising the pain to a particular area of the body. (cf. discussion of cases like these in §2.).

[a] As far as I am able to tell, in the entries given for KNOW in Evans et al (2004, p.443) only *bengkan* might be interpreted as a psycho-collocation, where √*beng* is a bound root that does not appear as an independent word which can be translated as 'mind'. Note that there is debate about whether √*beng* is independently meaningful (cf. discussion in Evans 2007, 2010, p.516; Goddard and Wierzbicka 2014, p.94-101).

[b] While I have not counted it here there is an example *yang-wonan* 'listen to' (Evans et al 2004:445) where *yang* 'language' is the psycho-noun. While language certainly has a relationship to the experiencer, this relationship does diverge from the relationship between the person-part and experiencer outlined in Section 2.1.2.

[c] "*Yolh-djare* [i.e. feeling-want, DM] was used occasionally with the sense 'like' (have affection for), i.e. instead of *djare*, which can be used alone." (Ponsonnet 2014b, p.443)

[d] There are no entries related to BE SAD in the English finder list in Evans et al (2004). However related terms can be found in other publications (e.g. Ponsonnet 2014b), though whether these can be understood as translatable to related terms in other languages is unclear.

[e] Based on DJONG compounds in Ponsonnet (2014b, p.217-218), where DJONG is interpreted as a psycho-noun but not a person-part. There are no entries related to FEAR in the English finder list in Evans et al (2004).

[f] Van Bik (1996, p.184) includes a construction of a form that appears to be similar to psycho-collocations as given in Van Bik (1998), where *rawl* 'diet, food, meal, refection' would be interpreted the psycho-noun. Note however that the interpretation of this construction as psycho-collocation only seems to make sense where for the sense of *rawl* as 'refection'.

[g] It might be the case that *lung* in *lungfahnak* 'pain' (Van Bik 1996, p.286) is a compound with the word *lung* 'liver, heart' and *fahnak* 'pain', though this remains speculative given the available evidence. See psycho-collocations with the same two words to describe someone feeling "bitter" in Van Bik (1998, p.212) and Yandt (2019, p.6). Note also that *lung* is the most frequent person-part for Hakha Lai entries in the database (cf. frequency in Figure 5.3).

In agreement with the analysis above, Table 5.3 suggests that psycho-collocation is typically used when referring to experiential situations in the emotion subdomain. Additionally, it suggests that there are differences in the kinds of meanings where psycho-collocation appears when looking beyond the emotion subdomain. Following from this, Table 5.3 suggests that there are differences within each of the subdomains that might be investigated further in future research. For example, while all languages except Hakha Lai appear to demonstrate psycho-collocation for either THINK and KNOW, it is not the case that each of these languages does so for *both* THINK and KNOW. This might be the result of language-specific differences in lexical semantics, which might be indicated by differences in the non-experiential meaning(s) of a given language's most frequent psycho-nouns. For example, when considering the perception subdomain, the fact that Mangap-Mbula is the only language which gave a plausible psycho-collocation as a definition for SEE may be linked to the productivity of *mata-* 'eye' across related constructions in Mangap-Mbula (cf. §5.2.1 below).

Table 5.3 also demonstrates two cases which also appear in the database and which are of importance for identifying the edges of the concept of psycho-collocation. These two edge cases are worth introducing here. Firstly, there are some cases in the database where the person-part which acts as the psycho-noun can be used to express an experiential

meaning without necessitating a psycho-mate (n=21). Table 5.3 above suggests an interesting property of examples of this sort as observed in this investigation: when considering “basic” experiential meanings, examples of this sort appear instead of psycho-collocation in the volition subdomain. The database demonstrates that examples of this sort are not limited to the volition subdomain. This is shown by the (non-exhaustive) examples in Table 5.4. Note that the database only contains entries of this type in Yucatec Maya and Mangap-Mbula.[53]

53 Though it should be said that these are the languages for which experiential constructions are most comprehensively surveyed, meaning the similar examples might appear for the other languages considered here, but did not appear in the surveyed publications (cf. §3.3 on database construction).

Table 5.4. Examples of person-parts which are used to express experiential meanings without requiring a psycho-mate.

Language	***Person-part***	***Example where person-part appears to be a predicate***
Mangap-Mbula[54]	*lele-* 'inside(s), feelings'	Y lele-POSS pa Z:Npersonal 'love, like someone' (Bugenhagen 2001, p.88; cf. Bugenhagen 2002, p.14)
	mata- 'eye'	mata-POSS pa 'look after, keep an eye on, watch over, care for' (Bugenhagen and Bugenhagen 2007, p.203)
	ni- 'being'	ni-POSS (mbura-) pa som 'not desiring' (Bugenhagen and Bugenhagen 2007, p.291)
	talŋa- 'ear'	talŋa-POSS som 'disobedient, unresponsive, not listening' (Bugenhagen and Bugenhagen 2007, p.505)
Yucatec Maya	*chi'* 'mouth' and *ni'* 'nose'	nip'chi'-t [taste:mouth-TRR] 'try, taste' (Verhoeven 2007, p.238) [55]
	ich 'eye'	ich-t [eye-TRR] 'fall in love, adore' (Verhoeven 2007, p.222)
	óol 'mind'	óot [mind:TRR] 'agree; wish, want' (Verhoeven 2007, p.222)

While examples like those in Table 5.4 are related to psycho-collocations in the sense that narrow psycho-collocation includes a person-part as the psycho-noun, they do not conform to the definition of broad psycho-

54 Mangap-Mbula also has an example with *kopo-* 'stomach' with the translation "take care of, work for, make sure all is well" (Bugenhagen and Bugenhagen 2007, p.131). This has not been included here given it is taken as not encoding an experiential situation.

55 This example is analysed differently in different sources: either as a transitivised noun compound in Bricker et al (1998, p.69), or as a person-part *chi'* 'mouth' which is incorporated into nip' 'taste' in Verhoeven (2007, p.238). Under Verhoeven's analysis this would not be considered one of the edge-cases discussed here. Examples like this, where two person-parts are present, are not discussed in detail here but appear several times in the database and would be worth investigating in future research.

collocation given in 2.3.2. This is because they only include one sign which is a non-grammatical lexeme (cf. Melčuk 2023, p.36-37). Note that some do appear alongside obligatory grammatical lexemes like the oblique marker *pa* and the negator *som* in Mangap-Mbula. Similarly, some appear with derivational morphemes like the transitiviser *-t* in Yucatec Maya.

A further edge-case shown in Table 5.3 above are those examples where a person-part is required for an experiential construction because of more general syntactic rules that deal with entire classes of lexical units, rather than because a specific person part necessitates a particular psycho mate in order to achieve the desired experiential meaning . For example, the Ewe expression *lãme vé ame* 'having pain in the body' conforms with the broad inclusion criteria given in §3.3.[56] However, the selection of *lãme* 'flesh inside' is not constrained by a speaker using *vé* 'pain' given *vé* can occur with a large number of words denoting parts of the body (cf. Awuku 2006, p.289-291). This means the example does not demonstrate constrained selection and would therefore not be a lexemic phraseme following Melčuk (2023, p.240), therefore not agreeing with

56 Note that here the Ewe word *ame* 'person' is used as a general term for 'person' in Ewe lexicographical tradition, and indicates where an argument denoting a person is required (cf. e.g. Awuku 2006, p.47). Additionally, note that expressions relating to bodily pain appear to be of this type for other languages in the sample also (e.g. Mangap-Mbula; Bugenhagen and Bugenhagen 2007, p.683).

the definition of (broad) psycho-collocation given in Section 2.3.2. As shown in Table 5.3, there are also examples of this type in Dalabon, Mangap-Mbula and Yucatec Maya in dictionary entries for FEEL PAIN (cf. Evans et al 2004, p.453; Bugenhagen and Bugenhagen 2007, p.683; Bricker et al 1998, p.310). I identified several entries of this type in the database, where an experiential construction includes a person-part and an experiential predicate, but where that predicate does not constrain the selection of the person-part (n=21). While the clearest cases were in the bodily sensation subdomain, it should be noted that there are likely more entries of the same type which I was not able to identify, particularly outside of the bodily sensation subdomain. Both cases can be understood as, related to, but beyond the bounds of psycho-collocation for the purposes of the revised definition in Section 5.4.

5.1.3 Beyond experiential subdomains in the database

Up to this point I have focussed on the overlap between and the subdomains of the experiential situation given in Verhoeven (2007; cf. §2.1.3) and demonstrated this for each language in Chapter 4. As noted in Section 2.2.2 research on psycho-collocation delineates subdomains which largely overlap with those given by Verhoeven, but they do not do so completely. This subsection briefly overviews ways in which database entries do not overlap with the experiential subdomains with an eye to the question of what constitutes a characterisation of the kinds of

meanings which are ‘typically’ expressed using psycho-collocations. In particular, I discuss three apparent divergences from the model given by Verhoeven (2007): psycho-collocations which denote social phenomena, which denote behaviours as opposed to feelings and which denote qualities or characteristics. I conclude by considering these three cases in light of the broader aim of characterising ‘typical’ psycho-collocation.

I begin with examples which both appear to agree with the definition of broad psycho-collocation and which refer to social phenomena of some sort. Note that Verhoeven discounts “verbs denoting the communication of an emotion” from the emotion subdomain given they are “thought of as representing action or communication concepts” rather than experiences (2007, p.46). Matisoff (1986, p.6) addressed cases like these by differentiating psycho-collocations with respect to whether they are more *subjective* or *inter-personal*. This is a distinct dimension of Matisoff’s classification where at one extreme subjective psycho-collocations are “inner-directed and self-contained” and at the other inter-personal psycho-collocations “denote mental states or attitudes that come into play primarily in connection with interpersonal relationships” (*ibid*). The significance of social aspects of experiential constructions has been discussed at length in relevant literature before and after Matisoff (1986), particularly on the language of emotion. The key example of this with respect to emotion is Wierzbicka (1999), who

describes emotion language as, for the most part, closely related to the social world of speakers. This point on emotion language is reiterated in later studies of experiential constructions (e.g. in Verhoeven 2007, p.44, Luraghi 2021, p.12). Some entries in the database suggest the role of the subjective/interpersonal dimension. First, note that there are a number of examples of this type which can also be understood as belonging to one or more subdomains of the experiential situation. Indicative examples are given as (42-43) below. Note that (43) appears to be both interpersonal in Matisoff's sense and outside of the emotion subdomain.

(42) *Dalabon*

nunh	wurdurd	mak	dja-marnu-yenjdju-ngiyan	delebon-walung [...].
DEM	child	NEG	3SG>2SG-BEN-talk-FUT	telephone-ABL

ka-h-dalu-yer-mu
3SG-R-mouth-shy.away-PR

'This child won't speak to you over the phone. He's too shy to talk [he's shy from the mouth]' (Ponsonnet 2014a, p.112; 20111207a_011_MT 48 [El])

(43)[57] *Mangap-Mbula*
Y mata-POSS imbol
(lit. Y eye is.strong)
'to be morally aware' (used for children from around the age of 8, who are then viewed as being responsible for doing wrong)
(Bugenhagen 2001, p.78)

Noting the role of this interpersonal dimension is significant for refining an initial characterisation of psycho-collocation. This is because doing so addresses the fact that there are constructions which are similar in form to other broad psycho-collocations in a given language but which do not neatly fall into an experiential subdomain. Indicative examples of these are given for each language in Table 5.5. For the most part, these are translated as English speech act verbs (e.g. "promise", "agree" etc.). Note that Mangap-Mbula examples of this sort were only found in the Mangap-Mbula dictionary (Bugenhagen and Bugenhagen 2007), rather than any publication on experiential constructions in particular.

57 Bugenhagen (2001, p.78) also gives the following definition for this expression in an adapted version of the Natural Semantic Metalanguage (cf. Goddard and Wierzbicka 2014 on NSM):

> Some people know that if a person does something it is bad
>
> Because of this, if someone does something like this people think:
>
> > "This person did something bad"
>
> Y is like this.

Table 5.5. Examples in the database which express interpersonal meanings

Language	*Example*
Dalabon	ngurlk-wurrhka HEART-surprise "Mislead with respect to conjugal relationship" (Ponsonnet 2014b, p.273)
Ewe	aβatsokãkã de ame ŋu lying on person flesh “To slander someone” (Awuku 2006, p.292)
Hakha Lai	ka luŋ ʔa tliŋ 1SG.POSS heart 3SG.SUBJ complete.I “I agree” (VanBik 1998, p.214, 228, 231)
Mangap-Mbula	Y mata-POSS ise ki Z Y eye ascend LOC Z “Look up to, respect, admire” (Bugenhagen 2001, p.81)[58]
Yucatec Maya	k'áat-óol-t wish:mind-TRR “Solicit, ask for, wish” (Verhoeven 2007, p.234)

Considering cases like these, which might be taken as representing psycho-collocations that cannot easily be sorted into an experiential subdomain given some person-parts are associated with an interpersonal dimension, a more comprehensive account of psycho-collocation across languages ought to carefully consider the role of this dimension. As noted in Osmond (2016, p.534) with respect to body part metaphor in Oceanic languages, this can hardly be done through sources like dictionaries and requires more detailed pragmatic and ethnographic

58 Noted as “to do with life and consciousness, the capacity to think and know” (Bugenhagen 2001, p.78).

information. While this restriction means the pragmatic and ethnographic detail cannot play a central role in this thesis, examples like those above are significant here because they suggest that considerations of these topics should play some role in research on psycho-collocation within and beyond mainland Southeast Asia.

Another example that Verhoeven excludes from their discussion of experiential constructions are “physiological and biochemical reactions (e.g., blush, get pale, increase of heart rate, etc.) [or] behavioural reactions (e.g., spitting, frowning, etc.)” (2007, p.44). Similarly, Verhoeven excludes “concepts of the expression of emotion such as ‘laugh’, ‘cry’ etc.” (Verhoeven 2007, p.46). In contrast, Matisoff (1986) does discuss “physical manifestations of emotion”, particularly in relation to person-parts on the more physical end of the psycho-somatic continuum mentioned in Section 5.2 (1986, p.10). However, meanings of this sort do not form a central part of his characterisation of (broad) psycho-collocation. Some analyses which apply the notion of psycho-collocations do refer to bodily reactions and behaviours, for example Lahaussois (In Press) and Bickel (1997, p.145), though they play a minor role in these cases. Nonetheless, expressions denoting associated behaviours and bodily reactions/processes have played a role in the description of experiential constructions, particularly relating to emotion language. For example, they play a role in Wierzbicka’s (1999) typology of emotion,

which features externally observable (involuntary) bodily processes/events and bodily sensations associated with, but not constitutive of, emotional experiences (cf. Ye 2002). Another example are the behavioural and somatic metonymies discussed in Ponsonnet and Luraghi's (2020) typology of emotion language in Australian languages which involves body parts. Meanings of this kind are instantiated in the database, with indicative examples given as Table 5.6 below. Note that Table 5.6 includes examples that are not in the emotion experiential subdomain.

Table 5.6. Examples in the database which denote physiological, biochemical or behavioural reactions rather than experiences themselves

Language	*Example*
Dalabon	dje-redjiredji face-laugh:REDUP “laughing” (Ponsonnet 2014b, p.191)
Ewe	é-ƒé ŋkú biā 3SG-POSS eye is.red "X's eye is red" (Ameka 2002, p.47); "To be envious of or covetous about an entity (persons or things)" (Ameka 2002, p.28); "... focussed all his attention" (Ameka 2002, p.48, quoting Hlomatsi 1994, p.54); Used to denote desperation in a phrase translating as "desperately look for" (Ameka 2002, p.48, quoting Ayeke, 1973, p.30)
Hakha Lai	ka ʔor ʔa fok 1SG.POSS throat 3SG.SUBJ protruded/erect.I “I want to eat” (VanBik 1998, p.209); “I am salivating (thinking about food)” (Yandt 2019, p.7)
Mangap-Mbula	kuli-POSS (i)sarakrak skin-POSS shivers “Have shivers up one's spine (in response to being cold or encountering the supernatural)” (Bugenhagen 2001, p.102); “Shiver with cold” (Bugenhagen and Bugenhagen 2007, p.142); “Have shivers up one's spine, hair “(Bugenhagen and Bugenhagen 2007, p.142); “Tremble with anger” (Bugenhagen and Bugenhagen 2007, p.142).
Yucatec Maya	ts'úuyni'-t bend:nose-TRR “Wrinkle nose (at bad odor)” (Verhoeven 2007, p.234)

Expressions like these, which refer to bodily reactions and behaviours associated with experiences, are significant for the construction of a refined definition of psycho-collocation for two reasons. Firstly, they are a problem case with respect to the characterisation of psycho-

collocations as a kind of phraseme given the argument that they demonstrate constrained selection is relatively unclear. I return to this point in Section 5.4. Secondly, they are a problem case with respect to the characterisation of psycho-collocations as experiential constructions given they might be understood as referring to bodily reactions and behaviours that are associated with experiences, but which do not constitute those experiences. This division is not entirely clear in the definitions collected here. For example, in the Hakha Lai example given in Table 5.6, the definition given by Van Bik (1998, p.209) does not suggest the example refers to a behaviour, though the definition given by Yandt's (2019, p.7) informant does. Considering cases like these I tentatively treat examples which are otherwise identical to psycho-collocations but which refer to bodily reactions and behaviours as a non-typical example of psycho-collocation in the revised definition given in Section 5.5.

Finally, I consider character traits and dispositions. As noted in Section 2.2.2, these are discussed in Matisoff's (1986) characterisation of the meanings expressed by (broad) psycho-collocation and cut across the subdomains of the experiential situation given in Verhoeven (2007). Character traits and dispositions have played a role in analyses which discuss the notion of psycho-collocation. For example, in Jaisser's (1990, p.164-167) analysis of *siab* 'liver' in White Hmong and Weissman's (2023,

p.435) study of *ini* 'heart' in Teposcolula Mixtec. Further, it is reasonable to understand some of the examples in the database considered here as character traits or dispositions. Noting that Matisoff (1986, p.5) gives GENEROUS, OPEN-MINDED and GULLIBLE as examples of character traits, Table 5.7 gives examples in the database which are understood as having similar meanings based on their translations. Note that Yucatec Maya's absence is likely because of exclusion rather than there being an absence of expressions of this sort in Yucatec Maya (cf. Verhoeven 2007, p.45-46). A difference between languages with respect to whether character traits and dispositions appear as psycho-collocations is also noted by Matisoff (1986, p.4), who notes that "Western languages rank higher on the scale of phenotypicality" with respect to psycho-collocations for traits and dispositions "so that they are more directly comparable with SE Asian languages".

Table 5.7. Examples in the database referring to character traits or dispositions

Language	*Example*
Dalabon	kangu-kerlk stomach-soft "Be a nice person" (Ponsonnet 2014b, p.406); "Be kind, generous, nice, friendly." (Ponsonnet 2014b, p.419)
Ewe	ɖɔme nyo stomach good ‘To be generous/kind’ (Awuku 2006, p.146)
Hakha Lai	ʔa luŋ ʔa raŋ 3SG.POSS heart 3SG.SUBJ quick.I 'He/She is impatient' (VanBik 1998, p.213); 'He makes judgements quickly' (Yandt 2019, p.9)
Mangap-Mbula	Y mata-POSS iŋgal itu-POSS Y eye pierces self ‘Y is careful and takes precautions’ (Bugenhagen 2001, p.79)

Similar to the other cases discussed in this section, traits and dispositions appear significant to a complete account of psycho-collocation but are beyond the information that appears across all of the lexicographical work drawn on here.

5.2 Comparison of person-parts in psycho-collocation

In this subsection I consider person-parts as the psycho-noun in psycho-collocations with respect to the experiential subdomains discussed in the previous sections. I first consider the uneven distribution of person-parts

in the database. I then investigate the relationship between person-parts and subdomains of the experiential situation with the aim of investigating whether any kind of generalisation can be made about person-parts in the context of psycho-collocation. I find three plausible groups of person-parts and note implications for a revised definition of psycho-collocation.

Note at the outset that there is no universal classification of person-parts across languages. This means, for example, that there is no reason to believe that the Dalabon word *ngerh-no*, Ewe *dzi*, Hakha Lai *luŋ* and Yucatec Maya *puksi'k'al* all have the same referent, even though they can each be translated into English as "heart" (cf. relevant definitions in Chapter 4). There are a few reasons for this. Firstly, research on the partonomy linguistic signs for person-parts has demonstrated that different languages divide the person along different lines (cf. Evans 2010, p.511-514). Further, terms for person-parts are prone to polysemy within and beyond the domain of person-parts (i.e. they allow multiple senses; cf. e.g. Kraska-Szlenk 2014; Liang, Xu & Ran 2024). This means that, even if it were the case that two different person-part words shared a physical referent, there is every chance that those words would not share their other senses. These two aspects of person-part words are illustrated in Sharifian's (2008) analysis of the Persian *del* 'heart-stomach' and English *heart*. The two words differ with respect to their

physical senses, where unlike *heart*, *del* "refers to the area of the body below the chest and above the pelvis, roughly similar to the area described in English by abdomen" (*ibid*, p.249). The words *del* and *heart* also differ when considering their non-physical senses where, for example, *del* is conceptualised as the "seat of thinking", closer to English 'mind' (*ibid*, p.255-256). Considering these differences between person-part words across languages, how do we attempt to compare psycho-collocations with respect to a psycho-noun which is a person-part?

One approach is by defining more abstract, higher level, categories of person-parts. It is certainly the case that authors have described groupings of person-parts within or across languages which might be employed here. For example, Matisoff (1986, p.9-10) notes that some psycho-nouns can be placed along a "psycho-somatic continuum", on which he distinguishes three types: mostly mental (e.g. English *mind*, *soul*, *spirit*), inherently ambiguous (e.g. English *heart*) and mostly physical (e.g. English *liver*, *guts*).[59] Another relevant categorisation is between external, internal, and abstract/psychological person-parts, which is applied, for example, by Ponsonnet (2014a) in the context of Dalabon emotion language and Vittrant (2013, p.266-267) in the context of Burmese psycho-collocations. But these approaches which employ

59 Although Matisoff (1986) discusses psycho-nouns in general in this section, it appears that these claims only refer to psycho-nouns which are person-parts, though note others like Bickel (1997, p.138-139) make similar distinctions between kinds of psycho-noun but also include abstract nouns (e.g. *desire*, *hunger* etc.).

more abstract groupings lead to a further question for our purposes: when considering the experiential subdomains associated with different person-parts, does the database imply a way of grouping person-parts in the context of psycho-collocation? In the remainder of this section I will investigate this question.

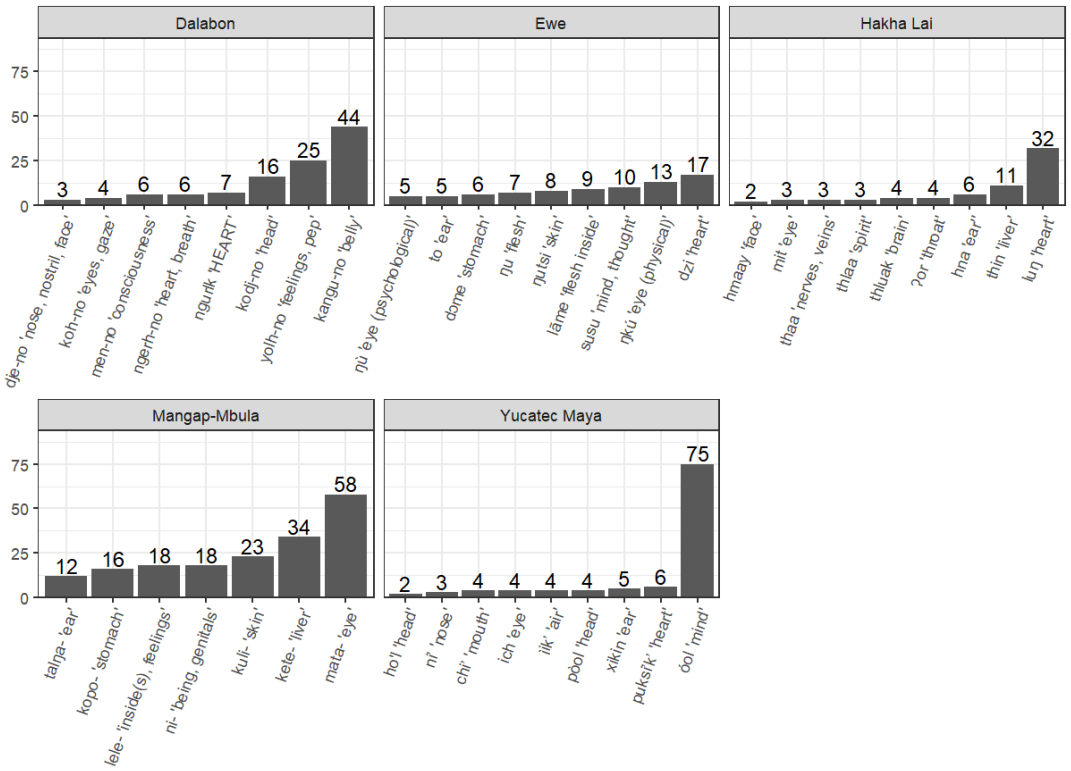
Firstly, note that in the database person-parts are not distributed evenly within each language. The difference between person-parts with respect to the number of unique pairings between psycho-nouns and psycho-mates is illustrated by Figure 5.3. This can be interpreted as some person-parts being more productive than others in the context of psycho-collocation.[60] For example, Figure 5.3 shows that for each language the most frequent person-part in the database occurs at least three times more often than the least frequent person-part.[61] Yucatec Maya is the clearest example of the uneven distribution of person-parts

60 This is not unreasonable when noting that the role of the frequency of psycho-nouns is well established in research which is of relevance to psycho-collocation. For example, McElhannon (1977, p.9) notes that "usually only a few [person parts] are very productive and some may occur only rarely [in body image idioms]". Similarly, Oey (1990, p.144-145) distinguishes "primary" and "secondary" psycho-nouns on the basis of both frequency and the kind of psycho-mate the psycho-noun is associated with (note §5.3 is of relevance to the second component of Oey's distinction).

61 There are five examples where two of the psycho-nouns considered here appear in the same database entry as the psycho-noun. For Figure 5.3 I've treated these cases as two different entries when counting person-parts.

in the database, where *óol* ‘mind’ appears far more than any other Yucatec Maya person part.

Figure 5.3. Number of person parts in unique mate-noun pairings by language



This uneven distribution is important to note because it is associated with differences in the association between person-parts and subdomains. This can be seen by computing Spearman’s rank correlation to assess the relationship between the frequency of a given person-part and number of subdomains recorded for entries with that person-part across languages (cf. Levshina 2015, Chapter 6). This test found that the frequency of a person-part was positively correlated with the number of subdomains appearing for that person-part, and that this

correlation is both moderately strong and statistically significant (Spearman's r = .65, df = 40, p < .001; with p-value based on large-sample approximation given tied ranks). Thus, the database considered here suggests that for person-parts which occur more often in psycho-collocations, those person-parts appear in a greater range of experiential subdomains.[62] Once again considering Yucatec Maya *óol* 'mind', this correlation is reflected in Verhoeven's characterisation of *óol* as a semantically bleached "experience marker", in contrast to other Yucatec Maya person-parts which appear in similar constructions (Verhoeven 2007, p.266, 345).[63] Looking more broadly to a revised definition of psycho-collocation, the above suggests that such a definition might note that typically person-parts which appear with a greater variety of psycho-mates are associated with a use across a greater range of experiential subdomains. Further research is needed to determine whether this generalisation is restricted to person-parts or applies to all psycho-nouns. For example, it seems less likely that a psycho-noun like

62 A few things should be noted here. Firstly, it is not entirely clear that the database satisfies the constraint that the sample is randomly drawn from the population, which is an assumption for checking the statistical significance of the correlation (Levshina 2015, p.132). Secondly, it is not the case that this correlation holds when considering every language individually: there is no statistically significant correlation between the frequency of a person-part and number of subdomains for Ewe or Hakha Lai when considered independently. I give statistics for each individual language as Appendix A.

63 Note that Yucatec Maya has no case marking (Verhoeven 2007, p.99), meaning that analysing *óol* as a case marker would disagree with other analyses of Yucatec Maya grammar. For an alternative conclusion, see Reh (1998) for an analysis of a similar phenomenon in the Lwo languages of eastern Africa which does conclude that the psycho-noun functions as a case marker.

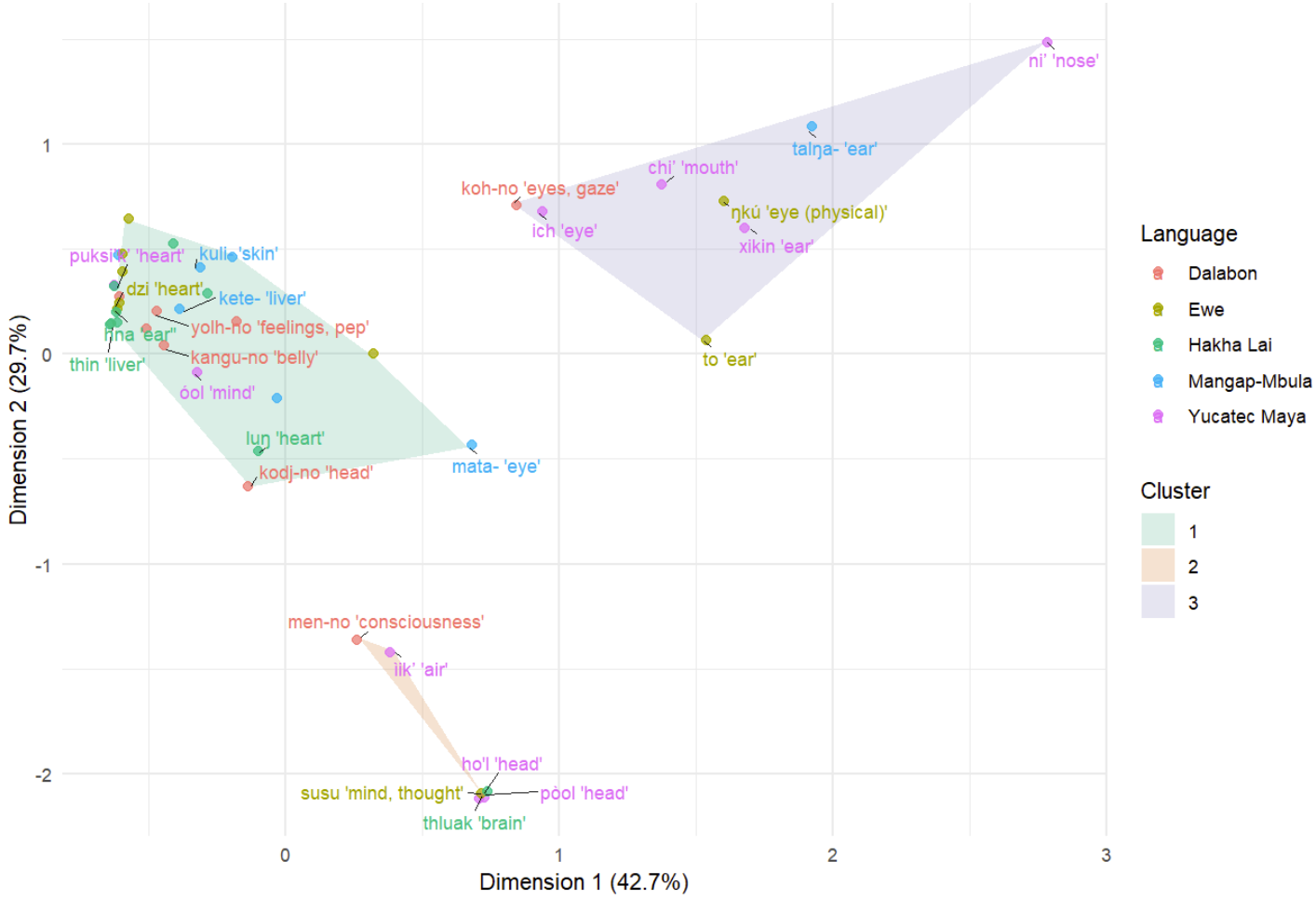
"happiness" would frequently appear outside of the emotion subdomain, regardless of the number of psycho-mates the word appears with.

Keeping in mind that the range of experiential subdomains associated with a person-part is correlated with its frequency, I now move to the question of deriving possible groupings of person-parts from associated experiential subdomains. To get a first picture of groupings between person-parts across languages in the database, I performed a correspondence analysis and generated hierarchical clustering from the results using the *FactoMineR* and *factoextra* packages in R (Kassambara & Mundt, 2020; Lê et al., 2008).[64] Figure 5.4 shows person-parts plotted on the first two (of five) dimensions following from the correspondence analysis. The x- and y- axes in Figure 5.4 represent the principal axes of association in the data, meaning they account for the majority of the variance between person-parts with respect to the experiential subdomains associated with that person-part in the database (~72.4%). The colours of individual points represent the language associated with a person-part, and the shaded polygons behind these points show the

64 Note that there is more statistical methodology which is relevant to the generation of this graph than I have scope to discuss here. Two points are worth mentioning here. Firstly, the method employed here is intended as exploratory analysis for the sake of hypothesis generation, rather than hypothesis testing. Secondly, while hierarchical clustering on principal components does order person-parts hierarchically, I have chosen not to investigate this hierarchy in detail here. For greater detail on the approach taken here and related methods though see discussion in Levshina (2015, Chapters 15, 18) and Holbert (2023).

three clusters returned by hierarchical clustering on correspondence analysis dimensions.

Figure 5.4. Grouping of person-parts across languages based on experiential subdomains of associated database entries[65]



Broadly, Figure 5.4 can be interpreted as suggesting a possible division of person-parts into three relatively homogeneous groups, where similarities within each group are larger than those between groups (cf.

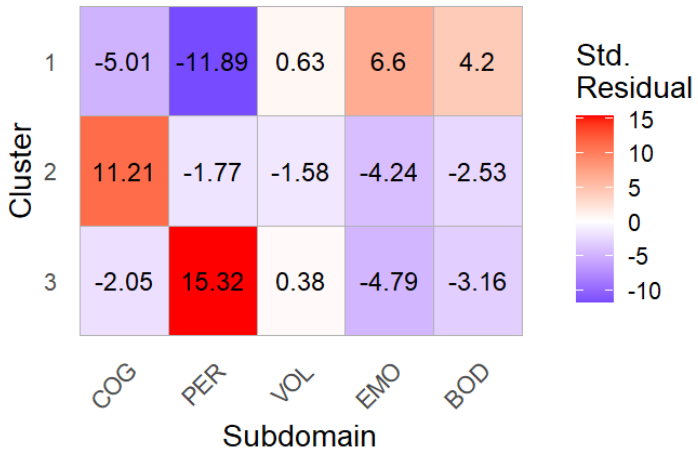
65 To reduce clutter in the plot (particularly in cluster 1) I have only included labels for the four most frequent person-parts in each language, as well as all person-parts in clusters 2 and 3.

Holbert, 2023; Levshina, 2015, p.369-373).[66] The degree of similarity is approximated by distance between points. When considering this distance, it should be noted that the two dimensions plotted in Figure 5.4 only capture a portion of the difference between points generated by the correspondence analysis. This portion is indicated by the percentage value on each axis of the Figure. The association between person-parts in these clusters and experiential subdomains can be investigated by considering the standardised residuals generated from a chi-squared test of association between the clusters in Figure 5.4, and the experiential subdomains associated with constituent person-parts.[67] These standardised residuals are given as Figure 5.5. This Figure highlights the over- or under-representation of each experiential subdomain within each cluster, where red cells (with positive values) indicate subdomains that are more frequent than expected and blue cells (with negative values) indicate subdomains which are less frequent than expected.

66 Note that the groupings discussed here are clearest when looking across languages, and that there might be more plausible groupings of person-parts within languages that do not necessarily line up perfectly with a derived cross-linguistic ideal. I've generated groupings when considering each language individually in Appendix B.

67 The contingency table used to generate Figures 5.4. and 5.5. is given as Appendix C.

Figure 5.5. Standardised residuals for chi-squared test of association between clusters given in Figure 5.4 and individual experiential subdomains



Considering positive residuals, Figure 5.5 suggests an association between cluster one and both the emotion and bodily sensation subdomains, cluster two and the cognition subdomain, and between cluster three and the perception subdomain. Given the relevant values in each row of Figure 5.5 are above 1.96 or below -1.96, Figure 5.5 can be interpreted as showing statistically significant deviations from the expected distribution of subdomains within each cluster (i.e. $p<0.05$; cf. Levshina 2015:221).[68] Additionally, there is no statistically significant

68 For the purpose of statistical significance the expected distribution is complete independence between cluster and subdomain. The expected distribution Positive or negative values show positive and negative association respectively. Note that it would not make sense to assume that the clusters are independent given they themselves are determined by association between person-parts and subdomains. However, the information in Figure 5.5 is useful to the extent that it makes the associations between particular subdomains and clusters in Figure 5.5 more explicit.

association between the volition subdomain and any of the three clusters given in Figure 5.4.

When interpreting Figure 5.5, it should be noted that the clusters given in Figure 5.4 are almost certainly influenced by the number of occurrences of a given person-part. As shown in Table 5.8, the number of unique mate-noun pairs in the database is lower for cluster two (n=32; 4.8%) and three (n=67; 10.1%) than it is for cluster one (n=567; 85.1%). This fact, alongside the demonstration of the relationship between frequency and range of experiential subdomains discussed above, suggests how the information in Figures 5.4 and 5.5 might be interpreted for the purpose of a refined definition of psycho-collocation. Such a definition would need to note that, in addition to the emotion subdomain being the most typical of psycho-collocation (cf. §5.1.1 and §5.1.2), the person-parts which appear in psycho-collocation can be grouped on the basis that they are associated with the less-typical cognition and perception subdomains.

Table 5.8. Frequency statistics for clusters in Figure 5.4

Cluster	***Number of unique mate-noun pairs (%)***	***Number of person-parts (%)***	***Average number of unique mate-noun pairs for a person-part in cluster***
1	567 (85.1%)	28 (66.7%)	20.25
2	32 (4.8%)	6 (14.3%)	5.3
3	67 (10.1%)	8 (19.0%)	8.375

There are some caveats that come along with this interpretation, the most important relating back to the comparability of person-part terms. It is probably the case that one can come up with a reason for why *susu* 'mind, thought' in cluster two relates to cognition, or why *to* 'ear' in cluster three relates to perception given their associated English translations. However, this should be done with caution, given there is apparent overlap in the meanings of person-parts in cluster one and those in either cluster two or three. For example, *póol* and *kodj-no* can both be translated as 'head', but appear in clusters one and two respectively. Similarly, *ich* and *mata-* can both be translated as 'eye' but appear in clusters one and three respectively. This is support for the point made at the beginning of this subsection: there are meaningful differences between apparently similar person-parts.[69] For the sake of future analysis of psycho-collocation it is therefore important to not assume a strong association between person-parts in two different languages which have been given apparently similar definitions.

To conclude, this section has proposed a possible grouping of person-parts in psycho-collocation with respect to experiential subdomain. I

69 It is worth noting that while Dalabon *kodj-no* 'head' and Mangap-Mbula *mata-* 'eye' appear in cluster one, they are placed more centrally in Figure 5.4, suggesting that they are more similar to person-parts in other clusters than may be the case for other person-parts in cluster one. This could suggest that there are similarities between person-parts with the same denotation but which only appear as weak signals in the database under investigation.

first noted that there are some person-parts which are more frequent than others in the database when considering unique pairings between psycho-nouns and psycho-mates. I then generated three groups of person-parts for the dataset by performing a correspondence analysis and generating hierarchical clustering from the results. Following this, I found that the most frequent person-parts in the dataset tend to be primarily associated with the emotion and bodily sensation subdomains. Some person-parts primarily appear in either the cognition or perception subdomain, though these person-parts tend to be less frequent when considering all unique pairings between psycho-nouns and psycho-mates. For the purpose of a revised definition of psycho-collocation, these suggest a more detailed characterisation that distinguishes typical and non-typical instances of psycho-collocation.

5.3 Psycho-collocations as lexemic phrasemes

In this section I consider whether all of the examples in the database are lexemic phrasemes. As introduced in Section 2.3.1, the two relevant parts of Melčuk's (2023, p.240) definition of lexemic phrasemes are that the selection of at least one component is constrained and that all of these components are lexemes. As a tool for doing so, I introduce Bickel's typology of "possessives of experience" in Section 5.3.1. In Section 5.3.2, I

then apply this typology in arguing that the majority of examples in the database demonstrate constrained selection before briefly justifying the claim that the psycho-mate and psycho-noun in each entry are always lexemes. In Section 5.3.3, I note how the languages under consideration differ with respect to Bickel's typology and address the question of the extent to which psycho-collocations can be understood as either idioms or collocations. This analysis concludes that for the most part there is a reasonable argument for psycho-collocations as lexemic phrasemes, meaning this aspect of the definition given in Section 2.3.1 can largely be retained in the revised preliminary definition given in Section 5.4.

5.3.1 Bickel's (1997) typology of the possessive of experience

In this section I consider Bickel's (1997) typology of the "possessives of experience" in Belhare, a Kiranti (Tibeto-Burman) language of Nepal.[70] Bickel's (1997) typology has been informative for studies of direct relevance to this thesis. For example, in Vittrant's (2013) characterisation of psycho-collocational expressives in Burmese and Lahaussois' (In Press) analysis of body part predicates in Thulung. As such it is

70 Bickel's (1997) study is ostensibly about "possessive of experience", though their application of Matisoff (1986) suggests a strong link between that concept and psycho-collocation. The application of Bickel (1997) in Vittrant (2013), which uses the term psycho-collocation, further suggests the similarity between the possessive of experience and psycho-collocation. See Bickel (2004) for an expansion of the possessive of experience into the "experiencer-as-possessive construction".

reasonable to consider the same typology here. I do so as a way of laying the foundation for Section 5.1.2., where I overview the extent to which entries in the database are lexical phrasemes by considering each of the types defined by Bickel (1997).

Bickel first distinguishes cases where the psycho-mate "occurs with only one noun and is 'morphanic' in Matisoff's sense" (Bickel 1997:141). A "morphan" is defined here as a sign which is "semantically empty but potentially related to full morphemes" (*ibid*).[71] For example, in Belhare the word *yuma* 'sleepy' only occurs alongside *mik* 'eye', as in (44).

(44) *Belhare*

a-**mik**	**yus**-e
1SG.POSS-**eye**	**sleepy**-PT

"I am sleepy" (Bickel 1997, p.141)

After distinguishing cases like (44) above, Bickel distinguishes five types of compositional psycho-collocation based on whether the psycho-noun

71 This is also referred to as an "eideme" in Bickel (1995, 1997), and a "degenerate lexeme" in Melčuk (2023). There is some variation here, the key one being that they are defined at different levels of specificity: where degenerate lexemes are defined as lexemes (Melčuk 2023, p.241) and eidemes as signs (Bickel 1997, p.141), and a type of "uniquely bound formative" in Matisoff (1986, p.46). The definition of morphan is made clear in Matisoff's (1973, p.60) grammar of Lahu: "By this term [(i.e. morphan)] we mean 'orphan morphs' that only occur in one or two compounds, and to which it is usually impossible to assign any meaning distinct from that of the compound as a whole." Note a key difference between morphan and both eideme and degenerate lexeme is that morphans are defined as being part of a compound. This aspect of the definition can be put aside to the extent that cross-linguistic studies of psycho-collocation since Matisoff do not always analyse psycho-collocations as compounds. An illustrative example of this comes from Vittrant, who notes that "psycho-collocations are a *semantic field* in which many MSEA languages *can* form distinctive compounds" (2019, p.662; emphasis mine).

and psycho-mate are specifically experiential in meaning.[72] This is shown in Table 5.9 (cf. also Verhoeven 2007, p.94-95).

Table 5.9. Bickel's (1997) typology of the lexical composition in "possessive of experience" constructions

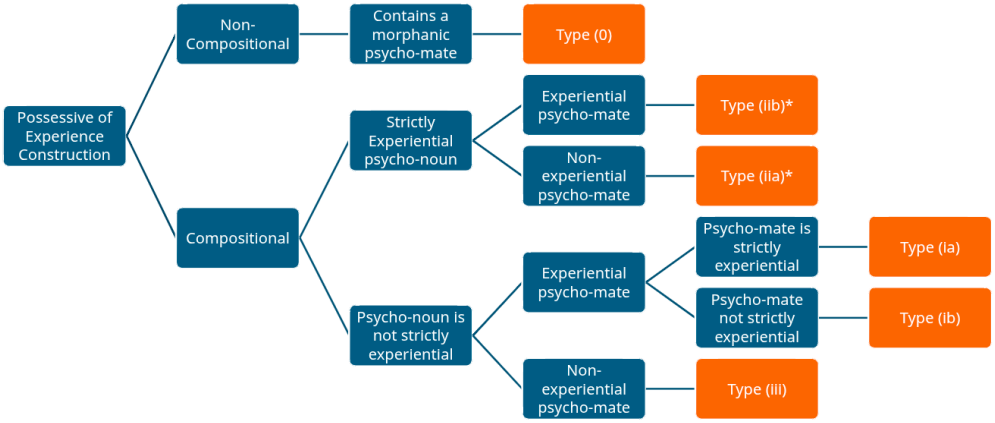
Type	*Excerpt characterising type*
(0)	"non-compositional collocations … The noun does not add information that is not already contained in the verb. The noun states, however, in a relation to a full-fledged lexical item" (Bickel 1997, p.141)
(ia)	"generalised noun and specialised 'psycho-mate', that is a verb subcategorised for a possessive of experience" (Bickel 1997, p.142)
(ib)	"The next type is similar to the preceding one except that the verb's experience meaning is part of a polysemy structure. As a consequence, the noun helps construct the experience function insofar as it disambiguates the verb" (Bickel 1997, p.142)
(iia)	"some nouns are strictly subcategorised for the experience function. In these cases, the verbal 'psycho-mate' has a general meaning" (Bickel 1997, p.144)
(iib)	"Experiential nouns also collocate with verbs that have themselves a strictly experiential meaning." (Bickel 1997, p.144)
(iii)	"Here, the experience meaning is an effect only of the possessive of experience construction itself. It is not pre-established by lexical semantics." (Bickel 1997, p.145)

Labelling non-compositional examples like (44) above as type (0), we can represent Bickel's typology in the terms used thus far as in Figure 5.1. In doing so, I take *specialised* to mean a psycho-noun or psycho-mate which encodes part of the experiential situation. *Generalised* then refers to a psycho-noun or psycho-mate which does not encode part of the

72 Note that Matisoff (1986, p.11-12) has an analogous idea of the "informativity" of the person-part with respect to the "metaphoricality" of the psycho-collocation. This aspect of psycho-collocation is investigated for some research in MSEA, for example in Van Bik (1998) with respect to Hakha Lai.

experiential situation (cf. §2.1.).[73] Note that while *NGURLK* 'heart' and *òol* 'mind' do not appear outside of experiential constructions, they are not understood as morphans here because they do not appear with a single psycho-mate.[74] I return to cases like these in Section 5.3.2. It should also be noted that types (iia) and (iib) include cases where the psycho-noun is an abstract noun denoting an experience, like "happiness", which are not considered here as a matter of scope.

Figure 5.6. Summary of the types of possessive of experience construction given in Bickel (1997)



*parts of Bickel's (1997) typology which are not considered in the scope of this thesis given they do not include a person-part as the psycho-noun (which is assumed to be non-experiential).

73 This distinction helps to clarify potential ambiguity in Matisoff's (1986, p.4) definition of the psycho-noun as "a noun with explicit psychological reference" which arises given Matisoff gives examples that do not appear to be exclusively experiential (e.g. Thai *nýa* 'flesh'; *ibid*, p.40).

74 These are not the only languages in the dataset which demonstrate psycho-nouns of this sort, though they are the only ones for which these psycho-nouns appear in the dataset. For example, *niŋ* in Hakha Lai (Proto-Tibeto-Burman **heart*; Van Bik 1998, p.217-218) would likely fall into this class but is not part of the dataset considered here.

5.3.2 Constrained selection of lexemic components

Following from the previous section, we can use Bickel's (1997) typology to disambiguate whether the mate-noun pairings in the database are examples of phrasemes.[75] As noted in Section 2.3.1., the relevant aspect of Melčuk's (2023, p.240) definition of *phraseme* is that they are constrained: broadly, to achieve some meaning (an experiential meaning in this case) the selection of one lexical unit necessarily requires the selection of another. In this section I first consider entries in the database where either the psycho-noun or psycho-mate can only appear in the psycho-collocation (type (0) above), followed by those where the predicate is non-experiential (type (iii) above), before considering those where the predicate takes an experiential meaning (types (ia) and (ib) above).[76] The argument given in this section allows for Section 5.1.3., which outlines the types given by Bickel (1997) with respect to each of the five languages under consideration and investigates the extent to

75 Noting that this argument excludes database entries where a person-part functions as a predicate and where a person-part is an oblique argument.

76 For the sake of completeness, it is worth briefly outlining why the identified psycho-noun and psycho-mate are identified as lexemes. At the outset, the psycho-noun and psycho-mate can be assumed to be lexemes where they appear as headwords in a dictionary given the assumption that headwords in dictionaries are always lexemes. While there were cases where I could find neither the psycho-noun nor the psycho-mate in a dictionary, I found no good reason to assume that the psycho-mate was not a lexeme in these cases. Considering the above I take it as true that both the identified psycho-noun and the psycho-mate in the database are lexemes.

which psycho-collocations are either idioms or collocations in Melčuk's (2023) general phraseology.

Consider first the entries where either the psycho-noun or psycho-mate can only appear in the psycho-collocation (type (0) above). The argument for these as phrasemes is fairly clear: if the morphan is selected, the other part of the psycho-collocation must be selected by definition because the morphan does not appear without that other component. Entries of type (0) are therefore phrasemes per Melčuk (2023). This is demonstrated in example (45) where, in parallel to (44) above, the Hakha Lai word *kuu* 'sleepy' can only be used in conjunction with *mit* 'eye' (Van Bik 2014, p.175).

(45) *Hakha Lai*

ka	**mit**	ʔa	**kuu**
1SG.POSS	**eye**	3SG.SUBJ	**sleepy.I**

"I am sleepy" (VanBik 1998, p.205, 231)

For entries in the database where the predicate is non-experiential (type (iii) above), there is strong evidence that that entry is a phraseme.[77] This is because the psycho-nouns in the database (which are always person-parts) do not have the meaning of the larger psycho-collocation independent of their psycho-mate.

This can be interpreted as the selection of a non-experiential psycho-mate requiring the selection of another lexical unit (the psycho-noun) to

77 This argument also applies to type (iia).

express an experiential meaning (i.e. that there is constrained selection). This means entries representing type (iii) are phrasemes per Melčuk (2023). A caveat to this argument is cases like those in Table 5.1. above where the person-part functions as the predicate. While it could be argued that these demonstrate that some person-parts have an experiential meaning and that they should therefore be understood as type (ii) above, their use outside of psycho-collocation suggests they do not have a strictly experiential meaning. Another caveat is Dalabon *NGURLK* 'heart', which as far as I am aware only occurs as an incorporated nominal in emotion compounds (cf. Ponsonnet 2014b, p.171-173).

Finally, consider the cases where the psycho-mate has an experiential meaning (types (i) above). These cases are less clear-cut than the others discussed in this section so far. This is because it is not always obvious whether a person-part is being selected because doing so is required by the psycho-mate, or whether the person-part is being selected because of more general syntactic rules dealing with entire classes of lexical units. For example, the Ewe expression *lãme vé ame* 'having pain in the body' conforms with the broad inclusion criteria given in Section 3.3. However, the selection of *lãme* 'flesh inside' is not constrained by a speaker using *vé* 'pain' given *vé* can occur with a large number of words denoting parts of the body (cf. Awuku 2006, p.289-291). I identified several entries of

this type, where an experiential construction includes a person-part and an experiential predicate, but where that predicate does not constrain the selection of the person-part (n=21; cf. §5.1.2). Note it is likely that there are more entries of the same type in the dataset which I was not able to identify, particularly outside of the bodily sensation subdomain.

We can refer to the division between types (ia) and (ib) for the remaining entries which include a psycho-mate with an experiential meaning. For entries where the psycho-mate is polysemous (type (ib) above), we can follow Bickel (1997, p.142) in arguing that the person-part disambiguates which sense of the psycho-mate is meant. Entries in type (ib) can be understood as phrasemes insofar as this is the case.[78] For example, in (46) below, *-saana* is glossed as "go bad" in Bugenhagen (1990, p.202). However, Bugenhagen and Bugenhagen (2007, p.446) define *-saana* as (among other things) "go bad", "decay", "tired", "worn out", "sick", and "feeble". Some of these senses are experiential, and others are not. Considering this, the occurrence of a person-part can be argued to be playing a role in differentiating the experiential sense of *-saana* from the non-experiential one. Entries like (46) are therefore understood as demonstrating the constrained selection of the person-part insofar as this is the case.[79]

78 Future research might consider whether this sense can be disambiguated by other means (context etc.), and if so what the function of the person-part is in these cases.

(46) *Mangap-Mbula*

ni-ŋ i-**saana**

body-1S.PSR 3S-**go.bad**

'I am exhausted.' (Bugenhagen 1990, p.202)

We now turn to the argument for constrained selection of the person part in database entries where the psycho-mate *only* allows an experiential meaning (type (ia) above). The argument that entries in the database of this type are phrasemes appears to be the weakest. Cases like these can be further divided. Firstly, here are a number of examples in the database where a psycho-mate only allows an experiential meaning, but that meaning is not in the same experiential subdomain as the larger psycho-collocation.[80] This is the case in (47).[81] Secondly, there are several cases where the psycho-mate only allows an experiential meaning and that meaning is in the same subdomain as the larger psycho-collocation, as in (48).

79 Note that it is not always clear that this is the case from the examples gathered and dictionary entries considered. It may turn out that database entries grouped under type (ib) here are in fact not phrasemes given the identified psycho-mate does not require the selection of the person-part for the meaning of the larger construction.

80 This is related to the "experiencer-of-response-to-emotions metonymies", "somatic metonymies" and "perception metaphors" identified in Ponsonnet and Luraghi's (2020) typology of the body in descriptions of emotion, which I noted in Section 5.1.3.

81 Noting that words related to pain were coded as part of the bodily sensation subdomain unless otherwise specified.

(47) *Yucatec Maya*
yah-óol-t
ache-mind-TRR
'be **depressed /sad** because of; **regret**' (Verhoeven 2007, p.222)

(48) *Dalabon*
koh-**wurrh(mu)**
eyes-**surprised/scared**
"Be surprised by the sight of something. *Koh-wurrh(mu)* can be used to refer to fright (minor), or simply to surprising information." (Ponsonnet 2014b, p.429; gloss derived from Ponsonnet 2014b, p.408)

Considering first cases like (47), although the meaning of the identified psycho-mate is experiential, the fact that the meaning of that psycho-mate is in a different experiential subdomain to the larger construction suggests that the person-part plays a disambiguating role similar to those identified for type (ib) above.[82] The selection of the person-part by the psycho-mate can therefore be argued for as was done for type (ib) above. Now considering cases like (48), there are some entries in the database where the occurrence of the person-part changes the meaning in a way that distinguishes the larger construction from the meaning of the psycho-mate alone, despite the psycho-mate and larger psycho-collocation being placed into different subdomains. This appears to be

82 This is an area where the subdomains are not sensitive enough, for example where the psycho-mate and the larger psycho-collocation are both in the emotion subdomain but where those two emotions seem distinct from one another based on the definitions given.

true for (48) above. These cases can similarly be identified as having constrained selection of the person-part on the grounds that the psycho-noun has a (comparatively subtle) disambiguating function with respect to the sense of the psycho-mate.

However there are some cases where a psycho-mate only has experiential sense(s) and is in the same subdomain as the larger psycho-collocation but where it is not clear that there is much difference at all between the meaning of the psycho-mate and the larger. For example, in (49), where the identified psycho-mate *k'oha'n* 'sick' combines with *óol* 'mind' to create a construction whose larger meaning is not ostensibly different to the meaning of the psycho-mate.

(49) *Yucatec Maya*
k'oha'n-kun POSS óol
sick-FACT POSS mind
'make feel **sickly, weak**' (Verhoeven 2007, p.252)

It may be that the issue here is the amount of lexicographical detail for the component parts of each database entry and/or that the psycho-mate in each of these cases can only appear alongside a person-part as in type (0) above. This is suggested by researchers considering psycho-collocations who question the sense that these psycho-nouns are semantically "redundant" from the perspective of their English translations (cf. e.g. Enfield, 2002; Jaisser, 1990, p.173-175; Matisoff, 1986, p.6-7; Sharma, 2015; *inter alia*). Keeping this in mind, it is nonetheless

unclear whether, in the database considered here, examples like (49) above, can be considered phrasemes. With respect to the definition given in Section 5.4, I tentatively take it as true that psycho-collocations can typically be interpreted as lexemic phrasemes.

This section has outlined an argument for entries in the database as phrasemes with respect to Bickel's (1997) typology of the possessive of experience (cf. §5.1.1). Based on the above argument it can be concluded that the majority of entries can plausibly be considered lexemic phrasemes following Melčuk (2023). This is because in each case there are plausible arguments that a particular person-part is required by the psycho-mate in order to achieve a particular meaning. These arguments are not of equal strength for each of Bickel's types however.

5.3.3 Proportion of Bickel type by language and whether psycho-collocations as idioms or collocations

In this section I consider the differences between the languages under investigation with respect to the types from Bickel (1997), which were discussed in the previous two sections. Considering only unique pairings between a psycho-noun and psycho-mate, Table 5.10 presents the distribution of database entries with respect to the types introduced in Section 5.3.1.[83] In this table I give the relevant percentage of unique

83 Note that it's not necessarily the case that looking at unique pairings between psycho-nouns and psycho-mates is the same as looking at unique psycho-collocations. This is because, for example, there are meaningful differences associated with the order of these two elements, with required stimulus arguments etc.

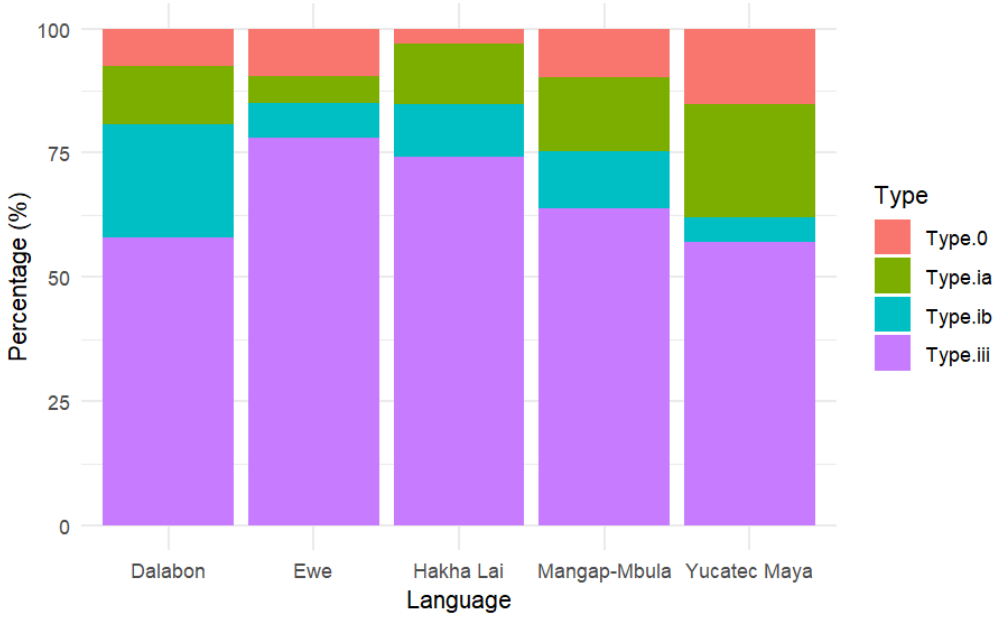
pairings within the language alongside a count of database entries. To more easily compare across languages, the percentage values given in Table 5.10 can be visualised as in Figure 5.2.[84]

Table 5.10. Unique noun-mate pairings for each language by types given in Bickel (1997)

Language	***Dalabon***		***Ewe***		***Hakha Lai***		***Mangap-Mbula***		***Yucatec Maya***	
Type	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
(0)	7	(7.53)	7	(9.59)	2	(3.03)	17	(9.77)	12	(15.19)
(ia)	11	(11.83)	4	(5.48)	8	(12.12)	26	(14.94)	18	(22.78)
(ib)	21	(22.58)	5	(6.85)	7	(10.60)	20	(11.49)	4	(5.06)
(iii)	54	(58.06)	57	(79.08)	49	(74.24)	111	(63.79)	45	(56.96)

84 Behaviours associated with emotions but not equivalent to them (e.g. crying) are considered non-experiential here.

Figure 5.7. Unique noun-mate pairings for each language by types given in Bickel (1997)



Type (iii), which include a psycho-mate without an experiential meaning, make up the majority of entries for each language in the database. Note however that the percentage of type (iii) entries ranges between ~56.96% for Yucatec Maya and ~79.08% for Hakha Lai. This suggests that across the languages sampled, the person-part most often plays a role in extending the meaning of the psycho-mate. This general claim groups together a great deal of variation between the meanings of the psycho-mates which have been grouped under type (iii) here.

Entries where the psycho-mate can have an experiential meaning (types (ia) and (ib)) make up the second largest portion of psycho-collocations for each language in the database. The sampled languages differ with respect to the proportion of types (ia) and (ib), the greatest difference between the two occurring for Dalabon and Yucatec Maya. Across

languages, this proportion is substantially smaller than the proportion of entries grouped as type (iii). Considering the argument that a database entry represents a lexemic phraseme is weakest for type (ia), the relatively low number of this type in Table 5.4 and Figure 5.2 suggests that the majority of database entries considered can comfortably be labelled lexemic phrasemes. It is worth noting again that it may be the case that some entries labelled as type (ia) have been misidentified, and could plausibly turn out to be type (0) on further analysis. Finally, for each language, type (0) makes up the smallest portion of entries overviewed in Table 5.4 and Figure 5.2.

5.3.4 (Non-)compositionality

In this section I briefly address (non-)compositionality as a way of addressing the question of whether psycho-collocations are idioms or collocations (introduced in §2.3.1). I conclude that while there is some evidence for the view that psycho-collocations can be either idioms or collocations, the evidence gathered here is insufficient to provide a definitive answer on the matter.

As noted in Section 2.3.1, while research related to psycho-collocation published since Matisoff (1986) often discusses psycho-collocations in relation to idioms, there are differences between these publications as to whether psycho-collocation is necessarily a kind of idiom. In Melčuk's (2023) view, a linguistic sign has three components: a signifier, a signified

and syntactics. A complex linguistic sign (i.e. a sign where its signifier includes more than one component) is non-compositional where that sign is not composed of the regular union of those components (Melčuk 2023, p.29). Melčuk gives the example of the lexeme REDNECK 'uneducated white man who lives and works in a rural area', where the signifieds of *red* 'the colour of blood' and *neck* 'the part of the body that joins the head to the shoulders' are not combined in a regular way to form "redneck"'. That is, the term does not refer to someone whose neck is red. In this case, the complex linguistic sign is non-compositional with respect to its signified, but non-compositionality also extends to cases where any of the signifier, signified or syntactics of the sign's components are not regularly combined (cf. Melčuk 2023, p.29).[85]

Looking now toward psycho-collocation, recall that on Melčuk's (2023) view, idioms and collocations are distinguished such that idioms are non-compositional and collocations are compositional (cf. §2.3.1). For the languages in the database, there is evidence that broad psycho-collocation can include both idioms and collocations. As an illustrative

85 This notion of regular union is determined by individual languages on Melčuk's view (2023, p.29). Note the distinction made by Melčuk between compositionality and transparency, where transparency is the ability of the Addressee to deduce the meaning of an expression from its components (Melčuk 2023, p.53). Transparency is not the focus of this thesis but might be investigated as an independent feature of psycho-collocation in future research. Ponsonnet points to a related distinction in their analysis of Dalabon emotion language when noting that for most compound lexemes in Dalabon, "while the lexicalized meaning is not compositional, speakers are aware of the sense of each morpheme" (2014, p.59).

example, we will take two cases from Dalabon, given as examples (50-51) below.

(50) *Dalabon*
kangu-yowyow(mu)
stomach-flow:REDUP
'be pleased, be nice' (Ponsonnet 2014a, p.108)

(51) *Dalabon*
koh-yer(mu)
eyes/gaze-shy.away
"avoid visual interactions for fear of others" (Ponsonnet 2014a, p.112; cf. Ponsonnet 2014b, p.223)

Examples (50-51) represent the distinction between incorporated nouns and lexicalised compounds in Dalabon. Ponsonnet describes (50) as part of a larger group of "lexified compounds" where *kangu-yowyow(mu)* is "clearly lexicalized given their non-compositional, non-predictable meaning" (Ponsonnet 2014a, p.105; cf. Ponsonnet 2014b, p.59, 222). We can take non-compositional meaning here as referring to non-compositionality with respect to the signified, as in Melčuk's REDNECK example above. Considering this, the class of lexicalised compounds in Dalabon can be understood as a class of idioms insofar as it is true that they demonstrate constrained selection. Ponsonnet contrasts lexicalised compounds with examples like (51), which are incorporated nouns in which an (absolutive) argument appears in the verb complex (2014a, p.104). In Dalabon these are regular and compositional (*ibid*). These

incorporated nouns can be understood as collocations in cases where they also demonstrate constrained selection (cf. §5.1.2). Additionally, both cases rely on the assumption that Ponsonnet's notion of compositionality is compatible with Melčuk's. Considering cases like (50-51) above for Dalabon, it is plausible that broad psycho-collocation encompasses both idioms and collocations. While it is not possible to characterise this relationship with any certainty given the database consists of examples from academic publications rather than primary sources (cf. e.g. discussion in Weissmann 2019, p.52), my investigation has suggested that the division between subdomains has the potential to be a fruitful area for future research.

5.4 A refined definition of psycho-collocation

This section gives a revised version of the definition given in Section 2.3.2 with respect to the discussion of psycho-collocation thus far. This definition should be taken as both refined and preliminary, as there are a number of areas of psycho-collocation which could not be investigated here. Note that in the below definition, some concepts which are defined in Chapter 2 are not defined in detail here. Instead, references to relevant subsections for each part of the definition appear in brackets. This is a development of the notion of "broad" psycho-collocation discussed in Section 2.1 and 2.3.2, where narrow psycho-collocation as discussed in Chapter 2. This definition leads into Chapter 6, where I discuss how psycho-collocation fits among concepts used which can be employed in comparing languages, limitations of this study, and potential directions for future research.

Psycho-collocation can be defined as follows, beginning with the notion of psycho-colocation in general before moving to aspects which might be understood as differentiating typical and atypical psycho-collocation.

The term ***psycho-collocation*** refers to a kind of **lexemic phraseme** (§2.3.1; §5.3.) which can be understood as **denoting an experiential situation** (§2.1.2; §5.1.1). **Psycho-collocations have two components:** the psycho-noun and the psycho-mate. The **psycho-noun** is defined as a

noun which has explicit psychological reference in the context of the psycho-collocation (§2.2.1; §5.3.1). Psycho-nouns can be divided into those which must always encode part of an experiential situation (e.g. English 'happiness') and those that do not necessarily need to do so (e.g. English 'heart') (§5.3.1). Person-parts are one class of psycho-noun in the latter category (§5.2). The **psycho-mate** is defined as any other lexeme(s) which must appear alongside the psycho-noun to achieve the intended meaning of the psycho-collocation (§2.2.1; §5.3.2).[86] Psycho-mates can be divided into those which can only appear with a psycho-noun, those which do not typically have an experiential meaning outside of psycho-collocation, and those which allow an experiential meaning when they appear outside of psycho-collocation (§5.3.1). The final category of psycho-mate can be further subdivided into those which only allow an experiential meaning and those which also allow non-experiential meanings (§5.3.1). While psycho-collocations appear to some extent across languages, in some languages they are essential for expressing some kinds of experiential meanings (S4.6; Chapter 4)

86 i.e. the selection of the psycho-mate is constrained by the psycho-noun (cf. Melčuk 2023, p.240; §2.3.1, §5.3). Note that I use the term *lexeme* to describe the psycho-mate where Matisoff (1986, p.4) used *morpheme*. The distinction being whether one considers grammatical morphemes as part of the psycho-mate. I have chosen not to do so here given that approach clashes with Melčuk's characterisation of the lexemic phraseme (Melčuk 2023, p.240, Chapter 3).

Psycho-collocations can appear across the experiential subdomains given in Verhoeven (2007; §5.1.1). To the extent that this is true, a typical psycho-collocation can be understood as a kind of experiential construction. Typically, psycho-collocations are used to denote experiential situations in the emotion subdomain, which can be understood as reflecting a more general tendency of experiential constructions as a whole (§5.1.1; §5.1.2). That the emotion subdomain is the most typical subdomain of psycho-collocation is demonstrated by both the number of unique pairs of psycho-mate and psycho-noun (§5.1.1), and the occurrence of psycho-collocation in relatively "basic" meanings across experiential subdomains (§5.1.2). In contrast to typical experiential constructions, psycho-collocations are often used for meanings which are less typical of experiential constructions as a whole, but might be understood are more typical of psycho-collocation (§5.1.3; cf. §2.2.2). These meanings include: speech acts and descriptions of social interactions; behaviours, physiological reactions or somatic reactions which are associated with, but not constitutive of, experiences; character traits or dispositions (§5.1.3).

There are some aspects of psycho-collocation which appear typical for cases where the psycho-noun is a person-part. On the basis of experiential subdomain, some groups of person-parts can be distinguished across languages given they are primarily associated with

one of either the perception or cognition experiential subdomain (§5.2). Person-parts in these two groups tend to be less frequent, with more frequent person-parts appearing across a wider range of experiential subdomains (§5.2). Person-parts outside of these groupings may all be grouped together, and are primarily associated with the emotion and volition experiential subdomains. Where the psycho-noun is a person-part, the psycho-mate typically does not have an experiential meaning beyond the psycho-collocation (§5.3.3).

6 General discussion

This chapter looks more broadly at the notion of psycho-collocation, and the applicability of this thesis to future studies which apply the term. In Section 6.1 I discuss how psycho-collocation fits among other concepts in comparative linguistics. In Section 6.2 I overview limitations of this study and considerations for future research in the area.

6.1 Psycho-collocation as a comparative concept

Psycho-collocation can be broadly characterised as one of a number of alternative “solutions” to the problem of denoting experiential situations (cf. Evans & Levinson, 2009, p.446 on linguistic problems and solutions). The content of this thesis can then be understood as investigating whether it is reasonable to say this solution recurs outside of mainland Southeast Asia, concluding that it is reasonable to do so. However, this leads to a further question about how psycho-collocation can be understood with respect to other terms used in comparative linguistics: to what extent is psycho-collocation the same sort of terminology as other terms in comparative linguistics? For example, does it make sense to refer to psycho-collocation in the same way one might refer to dative case marking, adjectives or predicative possession? In this section I address this question. Following Haspelmath (2010, 2025; cf. Croft, 2016)

I conclude that the notion of psycho-collocation is best understood as a *construction-strategy*, a formal means of expressing a *construction-function*, a comparative concept defined purely by some function (Haspelmath 2025, p.11). Further, I also conclude that experiential situations can be interpreted as a set of *construction-functions* and an experiential construction can be interpreted as any member of the set of construction-strategies which encodes an experiential situation (psycho-collocation being one such construction-strategy). Appealing to these higher-level terms is useful insofar as it allows general a characterisation of research on experiential constructions.

I begin with the notion of a *comparative concept*, defined by Haspelmath (2010) as

> "concepts created by comparative linguists for the specific purpose of crosslinguistic comparison. Unlike descriptive categories, they are not part of particular language systems and are not needed by descriptive linguistics of speakers. They are not psychologically real, and cannot be right or wrong" (Haspelmath 2010, p.665; cf. Haspelmath, 2018, 2025).

Comparative concepts are distinguished on this view from language-specific descriptive categories, which may be similar across language but are ultimately driven by description of individual languages (Haspelmath 2010, p.664). With the notion of a comparative concept in mind, some of the discussion in previous sections can be understood as

referring to comparative concepts. Section 2.2.1 can be understood as suggesting that psycho-collocation is treated as a comparative concept in research which primarily focusses on languages of mainland Southeast Asia. Further, Melčuk (2023) might be interpreted as an attempt to define comparative concepts in phraseology (e.g. *lexemic phraseme*, *idiom*, *collocation* etc.; cf. Melčuk 2023, p.3-4).[87] Finally, the definition given in Section 5.4 can be understood as first considering psycho-collocation as a comparative concept, before attempting to go beyond a unitary comparative concept when distinguishing more and less typical aspects of psycho-collocation per the database.

To further clarify the notion of *psycho-collocation*, as well as *experiential construction* and *experiential situation*, it is worth introducing a further distinction between two kinds of comparative concepts: construction-functions and construction-strategies (Haspelmath 2025).[88] In broad terms, construction-functions are comparative concepts which are

87 Verhoeven (2007) might be interpreted as being in part a consolidation of the comparative concepts used in research on experiential constructions, though this study is different to the extent that it is not entirely clear whether these concepts are treated as "psychologically real" in Haspelmath's sense. Neither Verhoeven (2007) nor Melčuk (2023) cast their projects explicitly in terms of comparative concepts. Further, it is worth noting that to my knowledge the notion of the comparative concept has not been explicitly applied to comparative studies in phraseology more generally.

88 I acknowledge that these terms are applied to comparative concepts in morphosyntax by Haspelmath (2025, following Croft 2016), and my focus has largely been lexical-semantic in this thesis. However, I have found the distinction also useful for the purposes of this section and so take it to be true that it can be extended to the topic of this section.

defined purely functionally, and construction-strategies are the “formal ways of expressing a particular meaning or function”, thereby being neither purely formal nor purely functional (*ibid*, p.11). For example, Haspelmath (2023) defines an adjective as a root which denotes a property concept. Here, *property* concept is a construction-function, and *adjective* is a corresponding construction-strategy. In the context of this thesis, an experiential construction can be thought of as any construction encoding an experiential situation (cf. §2.1.2). The *experiential construction* might therefore be understood as the set of construction-strategy in the sense that the term refers to all of the various formal ways of encoding the *experiential situation,* which would then be understood as a construction-function. There’s a difficulty with this distinction however: there are many kinds of experiential constructions in any given language, and there is no single mapping between experiential situations and experiential constructions across languages. This was demonstrated in Section 5.1.2, which showed that there are some (highly comparable) experiential meanings which may be encoded using psycho-collocation in some languages but which may not be encoded using psycho-collocation in others. Following from this, it may be more appropriate to think of the experiential situation as a set of different but related construction-functions, with the experiential construction remaining the set of the formal ways those construction-

functions might be expressed. Following from this, psycho-collocation can be understood as one possible formal means of expressing some part of the set of construction-functions referred to by the term *experiential situation*.

Understanding the relationship between experiential situations, experiential constructions, and psycho-collocation in this way reveals that psycho-collocation is somewhat different to other experiential constructions. As an example of another kind of experiential construction, consider the dative-experiencer construction, in which the use of a dative pronoun is conditioned by use of certain experiential predicates. This means, for example, that in Polish the experiencer needs to be marked with a dative pronoun for some experiential predicates, as in *zimno mi* [cold 1SG.DAT] 'I am cold' (cf. Haspelmath, 2001, p.66). Here the use of dative case is an example of a construction-strategy as defined above.[89] What is particularly notable about experiential constuctions is that their properties are often distinguished from "canonical" cases (cf. Aikhenvald et al., 2001)[90] For example, the dative-marked experiencer in the Polish dative-experiencer construction can be contrasted with the

89 I acknowledge that languages differ significantly with respect to what is labelled the dative case. However, as demonstrated in Haspelmath (2010, p.665-666), there is a plausible formulation of the dative case as a comparative concept. This is sufficient for the argument given here.

90 Note that a significant exception to this are affective case marking in the Nakh Daghestan languages of the Caucuses (cf. Comrie and Van den Berg 2006).

canonical subject, with respect to not only dative marking but also behavioural properties such as control of reflexive pronouns (Haspelmath 2001, p.71-72). The number of "non-canonical" cases relating to experiential meanings led Landau to note that "in just about any language where psych[ological] verbs have been studied in any depth, some special properties of these verbs have emerged" (Landau, 2010, p.6). Over the course of this investigation, it has been difficult to determine whether psycho-collocation can also be understood as "non-canonical" in the sense discussed above. A key problem is that it is not immediately obvious what psycho-collocation would be canonical or non-canonical relative to, as there is no clear analogue between psycho-collocation and a more general comparative concept in phraseology which is not centred on the description of experiential meanings. For example, while this thesis has found that psycho-collocation can reasonably be defined as a kind of lexemic phraseme, it is not the case that all lexemic phrasemes denoting experiential situations are psycho-collocations, given these phrasemes may not have a distinguishable psycho-noun and psycho-mate. This misalignment means it is not clear whether psycho-collocation can reasonably be analysed as a particularly non-canonical example of general phraseological comparative concepts like *idiom*, *collocation*, or *lexemic phraseme*. Notably, when looking within individual languages there are examples of the constructions

which have not been labelled psycho-collocation here. For example, Awuku (2006, p.140-141) describes physio-psychological concept expressions in Ewe as part of a larger class of “inherent complement verbs”. Future research would benefit from investigating this further, particularly with reference to the several distinguishing morphosyntactic features of psycho-collocation given for the languages surveyed in Chapter 4.[91]

6.2 Considerations and limitations for future investigations of psycho-collocation

In this section I overview considerations for future research on psycho-collocation. While the preceding chapters have noted limitations of this study where relevant, these considerations relate back to overarching limitations of this study just as much as they look forward to addressing those limitations in future studies. In particular, in this section I discuss the role of pragmatic markedness, of integrating anthropological and ethnographic information, of attempting to derive finer-grained divisions of the experiential situation, and of considering the structural relationships between the parts of the psycho-collocation. I conclude by

91 Note that several relevant publications make reference to “possessor raising” when discussing topics of direct relevance to psycho-collocation (Konnerth, 2014, p.337; Ponsonnet 2014a, p.70-71, 2015; cf. Verhoeven 2007, p.263). However, it appears that these authors link the phenomena they are observing to the term in different ways. Future research might investigate this further.

noting that current literature on Oceanic languages appears to go some way toward addressing these considerations, I identify these languages as a potential locus for future research on psycho-collocation. [92]

First, note the significance of pragmatic markedness for future research on psycho-collocation. Research of direct relevance to this study strongly suggests the importance of pragmatics. The exemplar of this is research on experiential constructions in African languages, where pragmatics have played a role in analysis of experiential constructions at least since work by Reh (e.g. Reh 1998; cf. Becher 2003; Zoch 2020). An example in Wolof is (52), for which Zoch notes that while the literal translation is pragmatically marked in English, the Wolof expression a "pragmatically rather neutral way to express anger (and therefore the appropriate translation in English is 'I am angry')" (2020, p.718). Pragmatic markedness has not been the focus of this thesis. As such, future research would go beyond the scope of this thesis by considering pragmatics.

92 Note there are some further considerations which are of potential relevance but are not discussed here. There are well-established questions concerning person-parts relating to the granularity of the semantic domain, as well as how it can be grouped and dissected (cf. §4.4; Evans 2010, p.511-515). Further, this study focusses on languages where psycho-collocation is a frequent or exclusive construction-strategy, but this is not a requirement of the definition of psycho-collocation (cf. §2.2.1).

(52) *Wolof (Atlantic-Congo; wolo1247)*

suma	xol	dafa	tang
my	heart	PFOC:3SG	be.hot

'I am angry.' (lit. 'My heart is hot') (Zoch 2020, p.715; Becher 2003, p.46)

The relationship between culture and psycho-collocation is highly relevant for research on the phenomena both within and across languages. Studies of phenomena relating to psycho-collocation frequently considers the effect of culture. For example, when overviewing "body part metaphors" in Oceanic languages Osmond (2016, p.527) notes the relationship between the throat and magical practice. Similarly, Ameka (2002) describes the Ewe cultural scripts underlying associations between *biã* 'redden, become red' and the person parts *ŋkú* "eye (physical)" and *ŋù* "eye (psychological)". A further example is Von Prince (2017, p.518), who notes a parallel between psycho-collocation and descriptions of meteorological events in Daakaka, a further Oceanic language of Vanuatu. Examples like these allow an enriched account of psycho-collocation on a few different fronts which together suggest that a future study of psycho-collocation would benefit from considering anthropological and ethnographical literature, as well as culturally oriented semantic research. For instance, these accounts would allow for a more comprehensive assessment of the compositionality of a psycho-mate, given dictionaries alone are insufficient for this purpose (cf. §5.3.4). Further, detailed information

about the culture associated with psycho-collocation could allow for a more informed understanding of whether the extension of person-parts into the experiential domain is “online”, or a conventionalised secondary meaning (Enfield 2002). This would add clarity to the division of psycho-collocations following Bickel (1997) in Section 5.3.1 (cf. Melčuk 2023, p.32-33). Additionally, this would allow more detailed comparisons of the metaphors and metonymies in psycho-collocations (cf. e.g. Ponsonnet and Luraginha 2020).

As well as the possibility of enriching existing lexicographical information by appealing to anthropological and ethnographical literature, it is worth noting the more general difficulty of trying to explicate experiential situations at all. Broadly, experiential meanings lack an ostensible denotation. That is, there is nothing “out there” which is clearly being referred to by the majority of experiential constructions. Instead, the closest one can get are indexes, such as physiological and biochemical reactions like blushing, or behavioural reactions like smiling (cf. §5.1.3). This somewhat separates experiential constructions from other topics in semantically oriented typology. For example, researchers creating a typology of colour terms can refer to physical aspects of visible light. The upshot of this are the fuzzy internal and external boundaries of the experiential situation which in turn led to the definition of *experiential situation* as a set of different construction-

functions given in Section 6.1. In this thesis I attempted to move through this fuzziness by appealing to a set of established subdomains of the experiential situation (cf. §2.1.3). However, as repeatedly noted in Chapter 5, these experiential subdomains are a relatively imprecise tool. While they have been useful here for establishing the general semantic range of psycho-collocation, future research would likely benefit from carefully considering a larger range of experiential meanings which have been established as relatively cross-translatable (cf. §5.1.2). However, given the general lack of ostensible denotation for experiential meanings, defining a cross-translatable set of these meanings would not be a simple task. A possible approach to deriving this set of meanings would be to consider the differences between the meanings associated with a given construction-strategy for both experiential and non-experiential meanings, and to note whether there are commonalities in the application of the construction-strategy. Performing this approach for all experiential constructions in one language, one might generate a semantic map (cf. Hartmann et al., 2016; Haspelmath, 2003) which represents which kinds of meanings are associated with which construction-strategies. Repeating this process over many languages might then demonstrate whether it is the case that there are a set of finer-grained divisions with respect to the construction-strategies used to encode experiential subdomains than those which were applied here.

Finally, future studies would benefit from focussing on the structural effects of different relationships between parts of the psycho-collocation. In their study of psycho-collocation in Burmese, Vittrant (2013) notes the relationship between the psycho-noun and the psycho-mate, and the relationship between the psycho-noun and the experiencer as being of particular importance. Different relationships between the psycho-noun and psycho-mate are illustrated by (53-54) below.[93] In (53), the psycho-noun *dzi* 'heart' and psycho-mate *dzɔ* 'straight' are part of a psycho-collocation denoting joy, whereas they are nominalised to form the abstract noun *dzidzɔ* 'happiness' in (54) (cf. Ameka 1990:157). What is particularly interesting about (54) is that the nominalised psycho-collocation *dzidzɔ* 'happiness' might itself be interpreted as a psycho-noun in a related but distinct psycho-collocation of its own, with *dzi* 'heart' appearing again as an oblique argument.[94]

93 There's an important distinction between the terms used for situation members in Section 2.1.2 and the use of the parts of the psycho-collocation here (cf. §2.2.1). This is because a psycho-noun is not necessarily a person-part, though person-parts are the kind of psycho-noun which has been the focus of this thesis. While it is not true that *dzi* 'heart' and *dzidzɔ* 'happiness' are both person-parts, they are both psycho-nouns following the definition developed in this thesis. Future research might attempt to systematise the differences between these two kinds of psycho-noun more thoroughly.

94 This phenomenon is not limited to Ewe. See, for example, Verhoeven (2007, p.267) on "double possessive constructions" in Yucatec Maya. Other apparent examples appear with *kangu-yirru(-no)* in Dalabon (Ponsonnet 2014, p.301-302). Note that in some of Chapter 5 I did not consider cases where the identified psycho-noun appears in an oblique argument. To reintegrate these cases, future research might consider more deeply whether examples like English *to keep in mind* are best interpreted as kinds of psycho-collocation. See Mulder (2021) for a clear discussion of psycho-collocation that divides the construction into four types based on the morphosyntactic relationship between psycho-noun and psycho-mate.

(53) *Ewe*
dzi dzɔ m
heart straight 1SG
“I was happy" (Ameka 1990, p.141);
“Heart straightened me. (I was happy.)" (Sevor 2015, p.97)

(54) *Ewe*
dzidzɔ yɔ́ mía-ɸé **dzi**-me
happiness fill 1PL-POSS heart-in
“We are full of joy” (Ameka 1990, p.163; Sevor 2015, p.100)

Different relationships between the psycho-noun and psycho-mate might be used as a way of structuring future studies on psycho-collocation. An example of this in a study of a single language is Mulder (2021), where four different types of psycho-collocation are delineated based on the morphosyntactic relationship between the psycho-noun and psycho-mate.[95] While these differences cannot be dwelt on here as a matter of scope, future research would also likely benefit from integrating these various relationships between the psycho-noun and psycho-mate into the larger picture of psycho-collocation as it occurs across languages. Noting now the relationship between the psycho-noun and the experiencer, the importance of this relationship is demonstrated by (55-56). These examples demonstrate two different ways the experiencer might appear overtly alongside the psycho-noun and psycho-mate,

95 These are: noun phrases, nominal clauses, verb phrases and light verb constructions.

where the experiencer is indexed on the verb in (55), and indexed with possessive marking as in (56). Future research would benefit from investigating variation between and within languages with respect to how this relationship is encoded in psycho-collocations.

(55) *Dalabon*
dja-h-men-yer-minj
2sg-R-ideas-shy.away-PP
"**you** were ashamed" (Ponsonnet 2014b, p.155; appears as part of larger example)

(56) *Yucatec Maya*

k'oha'n	**in**	w-óol
sick	**POSS.1.SG**	0-mind

'**I** feel weak/sick/weary.' (Verhoeven 2007, p.167; EMB_0529)

To conclude, this section has noted a number of considerations and limitations which might inform future research on psycho-collocation. In particular, this section focussed on the role of pragmatic markedness, of integrating anthropological and ethnographic information, of attempting to derive finer-grained divisions of the experiential situation, and of considering the structural relationships between the parts of the psycho-collocation. With these in mind, future research on psycho-collocation would likely benefit from investigating psycho-collocation as it occurs in the Oceanic languages. There are a few reasons for this that are directly related to the considerations and limitations given in this section. Those that are most relevant are elaborate systems of possessive

marking, as well as the existence of strong lexicographical, anthropological and ethnographical traditions associated with the language family (cf. Osmond 2016; Fox 2018 for overviews).[96] Most concretely, of the Oceanic language Mangap-mbula in this thesis can be taken as demonstrating of the applicability of the term *psycho-collocation* to an Oceanic language (cf. §4.4).

96 While Fox (2018) discusses Austronesian languages, claims contained therein apply to the Oceanic languages (a subfamily of Austronesian) more generally, and would be of use to future research.

7 Conclusion

This thesis has constituted an investigation of the term *psycho-collocation*, particularly with respect to the application of the term outside of mainland Southeast Asia and with an eye to potential terminological definitional, and conceptual issues. The clearest output of this investigation is a refined definition of psycho-collocation (cf. §5.4). This definition was constructed with respect to both the various uses of the term *psycho-collocation* in available research literature, and an investigation of how this term might be applied to a database of examples. The database investigated in this thesis represents a small but diverse sample of five languages: Dalabon, Ewe, Hakha Lai, Mangap-Mbula and Yucatec Maya. Chapter 4 considered what were identified as psycho-collocations in each of these languages individually, before Chapter 5 turned to a comparison across psycho-collocation in these languages. In doing so, a further output of this research are conclusions which constitute a first step to distinguishing "typical" psycho-collocation, as well as a number of informed provocations about future directions that inquiry in this area might take.

Beyond constructing a refined definition of psycho-collocation, this thesis is a contribution to the body of research on experiential constructions, which aims to consider all of the constructions used to

encode experiential situations in a given language. This thesis is a contribution in this context given psycho-collocation does not play a significant role in the majority of languages which have been the focus of past linguistic research relating to experiencers. As such, psycho-collocation has played a minor role in subsequent comparative study. This thesis works toward filling this gap. Doing so will allow researchers to better interrogate languages where psycho-collocation constitutes the dominant or exclusive means for expressing some kinds of experiences. While it is by no means certain that the notion of psycho-collocation is the best conceptual tool for doing so, this thesis nonetheless demonstrates that a picture of how human languages are used to convey what we are thinking, perceiving, feeling beings cannot be complete without an account of the data I investigate here. As such, this thesis is a contribution to a discussion of how this phenomena might be described and compared.

Reference List

Aikhenvald, A. Y., Dixon, R. M. W., & Onishi, M. (Eds.). (2001). *Non-canonical Marking of Subjects and Objects* (1st ed.). John Benjamins Publishing Company.

Aldai, G., & Wichmann, S. (2018). Statistical observations on hierarchies of transitivity. *Folia Linguistica, 52*(2), 249–281. https://doi.org/10.1515/flin-2018-0006

Ameka, F. K. (1990). The grammatical packaging of experiencers in Ewe: A study in the semantics of syntax. *Australian Journal of Linguistics, 10*(2), 139–181.

Ameka, F. K. (2002). Cultural scripting of body parts for emotions: On 'jealousy' and related emotions in Ewe. *Pragmatics & Cognition, 10*(1–2), 27–55.

Andrews, A. (2007). The major functions of the noun phrase. In T. Shopen (Ed.), *Language typology and syntactic description* (2nd ed., Vol. 1, pp. 132–222). Cambridge University Press.

Awuku, A. S. (2006). *Physio-psychological concept expressions in Ewe: A syntax-semantics lexicon* (Publication No. 304939714) [PhD dissertation, State University of New York at Buffalo]. ProQuest Dissertations & Theses. https://virtual.anu.edu.au/login/?url=https://www.proquest.com/dissertations-theses/physio-psychological-concept-expressions-ewe/docview/304939714/se-2?accountid=8330

Ayeke, K. (1973). *Hlɔ̃biabia [Taking revenge]*. Bureau of Ghana Languages.

Barth, D., & Schnell, S. (2021). *Understanding corpus linguistics*. Routledge.

Becher, J. (2003). Experiencer constructions in Wolof. *Hamburger Afrikanistische Arbeitspapiere, 2*, 1–89.

Bickel, B. (1997). The possessive of experience in Belhare. In D. Bradley (Ed.), *Tibeto-Burman languages of the Himalayas* (pp. 135–155). Pacific Linguistics.

Bickel, B. (2000). On the syntax of agreement in Tibeto-Burman. *Studies in Language, 24*(3), 583–610.

Bickel, B. (2004). The syntax of experiencers in the Himalayas. In P. Bhaskararao & K. V. Subbarao (Eds.), *Non-nominative subjects* (Vol. 1, pp. 77–112). John Benjamins Publishing Company.

Bossong, G. (1998). Le marquage de l'expérient dans les langues de l'Europe. In J. Feuillet (Ed.), *Actance et valence dans les langues de l'Europe* (pp. 259–294). De Gruyter Mouton. https://doi.org/10.1515/9783110804485.259

Bricker, V. R., Po'ot Yah, E., & de Po'ot, O. D. (1998). *A dictionary of the Maya language as spoken in Hocabá, Yucatán*. University of Utah Press.

Bugenhagen, R. D. (1990). Experiential constructions in Mangap-Mbula. *Australian Journal of Linguistics, 10*(2), 183–215. https://doi.org/10.1080/07268609008599441

Bugenhagen, R. D. (1995). *A grammar of Mangap-Mbula: An Austronesian language of Papua New Guinea*. Pacific Linguistics.

Bugenhagen, R. D. (2001). Emotions and the nature of persons in Mbula. In J. Harkins & A. Wierzbicka (Eds.), *Emotions in crosslinguistic perspective* (pp. 69–114). De Gruyter.

Bugenhagen, R. D. (2002). The syntax of semantic primes in Mangaaba-Mbula. In A. Wierzbicka & C. Goddard (Eds.), *Meaning and universal grammar: Theory and empirical findings* (Vol. 2, pp. 1–64). John Benjamins Publishing Company. https://doi.org/10.1075/slcs.61.06bug

Bugenhagen, S. E., & Bugenhagen, R. D. (2007). *Ro ta ipiyooto sua mbula uunu = Mbula-English dictionary*. Summer Institute of Linguistics.

Carnesale, L. (2023). Experiential constructions in Hindi: The encoding of perceptions and bodily sensations. *Journal of South Asian Languages and Linguistics, 10*(2), 155–187. https://doi.org/10.1515/jsall-2024-2013

Clark, M. (1989). Hmong and areal South-East Asia. *Pacific Linguistics. Series A, 80*, 175–230.

Comrie, B., & van den Berg, H. (2006). Experiencer constructions in Daghestanian languages. In I. Bornkessel, M. Schlesewsky, B. Comrie, & A. D. Friederici (Eds.), *Semantic role universals and argument linking: Theoretical, typological and psycholinguistic perspectives* (pp. 127–154). De Gruyter. https://doi.org/10.1515/9783110219272.127

Croft, W. (2002). *Typology and universals* (2nd ed.). Cambridge University Press.

Croft, W. (2016). Comparative concepts and language-specific categories: Theory and practice. *Linguistic Typology, 20*(2), 377–393. https://doi.org/10.1515/lingty-2016-0012

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge University Press.

Dik, S. C. (1981). *Functional grammar*. De Gruyter. https://doi.org/10.1515/9783112420126

Enfield, N. J. (2002). Semantic analysis of body parts in emotion terminology: Avoiding the exoticisms of "obstinate monosemy" and "online extension." *Pragmatics & Cognition, 10*(1–2), 85–106.

Enfield, N. J. (2021). *The languages of mainland Southeast Asia*. Cambridge University Press.

Evans, N. (2004). Experiencer objects in Iwaidjan languages (Australia). In P. Bhaskararao & K. V. Subbarao (Eds.), *Non-nominative subjects* (Vol. 1, pp. 169–192). John Benjamins Publishing Company. https://doi.org/10.1075/tsl.60.10eva

Evans, N. (2007). Standing up your mind: Remembering in Dalabon. In M. Amberber (Ed.), *The language of memory in a crosslinguistic perspective* (pp. 67–95). John Benjamins Publishing Company.

Evans, N. (2010). Semantic typology. In J. J. Song (Ed.), *The Oxford handbook of linguistic typology* (1st ed., pp. 504–533). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199281251.013.0024

Evans, N. (2021). Social cognition in Dalabon. *Language Documentation & Conservation, 12*, 22–84.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*(5), 429–448.

Evans, N., Merlan, F., & Tukumba, M. (2004). *A first dictionary of Dalabon (Ngalkbon)*. Maningrida Arts and Culture.

Fedriani, C. (2014). *Experiential constructions in Latin*. Brill.

Fox, J. (2018). Towards a comparative study of the structure of sentiment among the Austronesians. In J. Fox (Ed.), *Expressions of Austronesian thought and emotions* (pp. 129–148). ANU Press.

Goddard, C., & Wierzbicka, A. (2014). Semantic fieldwork and lexical universals. *Studies in Language, 38*(1), 80–127.

Grimshaw, J. B. (1990). *Argument structure*. MIT Press.

Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2024). *Glottolog 5.1* (Version 5.1). Max Planck Institute for Evolutionary Anthropology. (Available online at glottolog.org; accessed 2024-11-29.)

Hartmann, I., Haspelmath, M., & Cysouw, M. (2016). Identifying semantic role clusters and alignment types via microrole coexpression tendencies. In *[edited volume]* (pp. 27–49). https://doi.org/10.1075/bct.88.02har

Hartmann, I., Haspelmath, M., & Taylor, B. (Eds.). (2013). *The Valency Patterns Leipzig online database*. Max Planck Institute for Evolutionary Anthropology. https://valpal.info/

Haspelmath, M. (2001). Non-canonically marking of core arguments in European languages. In A. Y. Aikhenvald, R. M. W. Dixon, & M. Onishi (Eds.), *Non-canonical marking of subjects and objects* (pp. 53–83). John Benjamins Publishing Company.

Haspelmath, M. (2003). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In M. Tomasello (Ed.), *The new psychology of language* (Vol. 2, pp. 211–243). Lawrence Erlbaum Associates.

Haspelmath, M. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language, 86*(3), 663–687.

Haspelmath, M. (2018). How comparative concepts and descriptive linguistic categories are different. In D. Olmen, T. Mortelmans, & F. Brisard (Eds.), *Aspects of linguistic variation* (pp. 83–114). De Gruyter Mouton Berlin.

Haspelmath, M. (2023). Word-Class Universals and Language-Particular Analysis. In E. Van Lier (Ed.), *The Oxford Handbook of Word Classes* (1st ed., pp. 15–40). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198852889.013.2

Haspelmath, M. (2025). *Construction-functions versus construction-strategies*. Language Science Press. https://doi.org/10.5281/ZENODO.16838056

Holbert, C. (2023, August 20). *Clustering on principal component analysis*. Charles Holbert. https://www.cfholbert.com/blog/cluster-pca/

Jaisser, A. (1990). Delivering an introduction to psycho-collocations with SIAB in White Hmong. *Linguistics of the Tibeto-Burman Area*.

Kassambara, A., & Mundt, F. (2020). *factoextra: Extract and visualize the results of multivariate data analyses* [R package manual]. https://CRAN.R-project.org/package=factoextra

Kimoto, Y., Shiohara, A., Barth, D., Evans, N., Kogura, N., Arka, I. W., Pratiwi, D. P. E., Kasuga, Y., Kawakami, C., & Kurabe, K. (2024). Syntactic embedding or parataxis?: Corpus-based typology of complementation in language use. *Language Documentation & Conservation [Special Publication]*, *12*, 126–162.

Konnerth, L. A. (2014). *A grammar of Karbi* (Publication No. 1535090744) [PhD dissertation, University of Oregon]. ProQuest Dissertations & Theses Global. https://www.proquest.com/pqdtglobal/docview/1535090744/

Kraska-Szlenk, I. (2014). Semantic extensions of body part terms: Common patterns and their interpretation. *Language Sciences*, *44*, 15–39. https://doi.org/10.1016/j.langsci.2014.02.002

Lahaussois, A. (In Press). Thulung body part predicates and the expression of emotional and mental states. *South Asian Linguistics Studies: Recent Advances and Current Trends*.

Landau, I. (2010). *The locative syntax of experiencers*. MIT Press.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: A package for multivariate analysis. *Journal of Statistical Software, 25*(1), 1–18. https://doi.org/10.18637/jss.v025.i01

Lehmann, C. (1985). On grammatical relationality. *Folia Linguistica, 19*.

Lehmann, C. (2002). *Possession in Yucatec Maya* (2nd ed.). Lincom Europa. https://www.christianlehmann.eu/publ/ASSidUE10.pdf

Lehmann, C. (2015). Valency classes in Yucatec Maya. In B. Comrie & A. Malchukov (Eds.), *Case studies from Austronesia, the Pacific, the Americas, and theoretical outlook* (Vol. 2, pp. 1427–1480). De Gruyter. https://doi.org/10.1515/9783110429343-014

Lehmann, C. (2022). Foundations of body-part grammar. In R. Zariquiey & P. M. Valenzuela (Eds.), *The grammar of body-part expressions: A view from the Americas* (pp. 14–76). Oxford University Press.

Lehmann, C., Shin, Y.-M., & Verhoeven, E. (2000). Unfolding of situation perspectives as a typological characteristic of languages. *STUF – Language Typology and Universals, 53*(1). https://doi.org/10.1524/stuf.2000.53.1.71

Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. John Benjamins. https://doi.org/10.1075/z.195

Liang, Y., Xu, K., & Ran, Q. (2024). Shared structure of fundamental human experience revealed by polysemy network of basic vocabularies across languages. *Scientific Reports*, *14*(1), 5877. https://doi.org/10.1038/s41598-024-56571-8

Lichtenberk, F. (1985). Possessive constructions in Oceanic languages and Proto-Oceanic. In A. Pawley & C. Louis (Eds.), *Austronesian linguistics at the 15th Pacific Science Congress* (pp. 93–140).

Luraghi, S. (2021). *Experiential verbs in Homeric Greek: A constructional approach*. Brill.

Malchukov, A. (2015). Valency classes and alternations: Parameters of variation. In A. Malchukov & B. Comrie (Eds.), *Valency classes in the world's languages* (Vol. 1, pp. 73–130). De Gruyter Mouton.

Matisoff, J. (1986). Hearts and minds in Southeast Asian languages and English: An essay in the comparative lexical semantics of psycho-collocations. *Cahiers de Linguistique Asie Orientale, 15*(1), 5–57.

McElhanon, K. A. (1977). Body image idioms in Irianese and Papua New Guinean languages. *Irian, 6*(3), 3–27.

Melčuk, I. A. (2023). *General phraseology: Theory and practice* (Vol. 36, 1st ed.). John Benjamins. https://doi.org/10.1075/lis.36

Miestamo, M., Bakker, D., & Arppe, A. (2016). Sampling for variety. *Linguistic Typology, 20*(2), 233–296.

Miłosz, C. (1986). *Separate notebooks* (A. Miłosz, Trans.). Ecco Press.

Mosel, U. (2006). Grammaticography: The art and craft of writing grammars. In *Catching language* (pp. 41–68). De Gruyter.

Mulder, M. (2021). Morphosyntax of Muklom psycho-collocations. In *Proceedings of the 11th Payap University and UNRN Research Symposium*. UNRN Research Symposium.

Musgrave, S. (2006). Complex Emotion Predicates in Eastern Indonesia: Evidence for Language Contact? In Y. Matras, A. McMahon, & N. Vincent (Eds.), *Linguistic Areas: Convergence in Historical and Typological Perspective* (pp. 227–243). Palgrave Macmillan UK. https://doi.org/10.1057/9780230287617_9

Næss, Å. (2007). *Prototypical transitivity*. John Benjamins Publishing Company.

Neuwirth, E. (2022). *RColorBrewer: ColorBrewer palettes* [R package manual]. https://CRAN.R-project.org/package=RColorBrewer

Newman, J. (2009). *The linguistics of eating and drinking*. John Benjamins.

Oey, E. M. (1990). Psycho-collocations in Malay: A Southeast Asian areal feature. *Linguistics of the Tibeto-Burman Area*.

Osmond, M. (2016). Body part metaphors. In M. Ross, A. Pawley, & M. Osmond (Eds.), *The lexicon of Proto-Oceanic: The culture and environment of ancestral Oceanic society* (Vol. 5, pp. 519–534). Pacific Linguistics.

Peterson, D. A. (1998). The morphosyntax of transitivization in Lai (Haka Chin). *Linguistics of the Tibeto-Burman Area*, *21*(1), 87–153. https://dx.doi.org/10.32655/LTBA.21.1.03

Peterson, D. A. (2014). Hakha Lai. In *The Sino-Tibetan languages* (2nd ed., pp. 258–276). Routledge.

Ponsonnet, M. (2009). Aspects of the semantics of intellectual subjectivity in Dalabon (south-western Arnhem Land). *Australian Aboriginal Studies, 1*, 16–28.

Ponsonnet, M. (2010). Aspects of the semantics of emotions and feelings in Dalabon. *The Australian Journal of Anthropology, 21*(3), 367–389.

Ponsonnet, M. (2014a). Figurative and non-figurative use of body-part words in descriptions of emotions in Dalabon (Northern Australia). *International Journal of Language and Culture, 1*(1), 98–130.

Ponsonnet, M. (2014b). *The language of emotions: The case of Dalabon (Australia)* (Vol. 4). John Benjamins Publishing Company.

Ponsonnet, M. (2015). Nominal subclasses in Dalabon (South-Western Arnhem Land). *Australian Journal of Linguistics, 35*(1), 1–52. https://doi.org/10.1080/07268602.2015.976900

Ponsonnet, M., & Laginha, K.-J. (2020). The role of the body in descriptions of emotions: A typology of the Australian continent. *Pragmatics & Cognition, 27*(1), 20–82.

Reh, M. (1998). The role of metaphor in grammar: Experiencer constructions. In M. Reh (Ed.), *Experiens-Kodierung in afrikanischen Sprachen typologisch gesehen* (pp. 25–40). Universität Hamburg.

Rijkhoff, J., Bakker, D., Hengeveld, K., & Kahrel, P. (1993). A method of language sampling. *Studies in Language, 17*(1), 169–203. https://doi.org/10.1075/sl.17.1.07rij

Rongier, J. (1995). *Dictionnaire français–ewé, suivi d'un index ewé–français* (p. 560). Éditions Karthala & Agence de Coopération Culturelle et Technique.

Sevor, G. (2015). *A linguistic analysis of Ewe metaphorical expressions* [MPhil thesis]. University of Ghana.

Sharifian, F. (2008). *Culture, body, and language: Conceptualizations of internal body organs across cultures and languages* (1st ed.). Mouton de Gruyter. https://doi.org/10.1515/9783110199109

Sharma, H. S. (2015). Manipuri psycho-collocations. *International Journal of Dravidian Linguistics, 44*(2), 105–114.

Simon, C., & Reh, M. (1999). Experienskonstruktionen in afrikanischen Sprachen. *Sprachtypologie und Universalienforschung, 52*, 167–182.

Slowikowski, K. (2024). *ggrepel: Automatically position non-overlapping text labels with ggplot2* [R package manual]. https://CRAN.R-project.org/package=ggrepel

Tsunoda, T. (1981). Split case-marking patterns in verb-types and tense/aspect/mood. *Linguistics, 19*(5–6). https://doi.org/10.1515/ling.1981.19.5-6.389

Tsunoda, T. (1985). Remarks on transitivity. *Journal of Linguistics, 21*(2), 385–396.

Van Bik, D. (1996). *English–Chin (Hakha) dictionary* (2nd ed.). Chin Association of Christian Communication.

Van Bik, D. (2014). *Chin (Laiholh) to English dictionary*. https://digital.library.unt.edu/ark:/67531/metadc1872869/

Van Bik, K. (2023). Issues of lexicon in South-Central Tibeto-Burman (Kuki-Chin). *Himalayan Linguistics, 22*(1).

Van-Bik, K. (1998). Lai psycho-collocation. *Linguistics of the Tibeto-Burman Area*.

VanBik, K. (2010). The syntax of psycho-collocations in Hakha Lai. *Linguistics of the Tibeto-Burman Area, 33*(2), 137–150.

Verhoeven, E. (2007). *Experiential constructions in Yucatec Maya: A typologically based analysis of a functional domain in a Mayan language* (Vol. 87). John Benjamins Publishing Company.

Vittrant, A. (2013). Psycho-collocational expressives in Burmese. In *The Aesthetics of Grammar: Sound and Meaning in the Languages of Mainland Southeast Asia* (pp. 255–279). Cambridge University Press.

Vittrant, A., & Watkins, J. (2019). Guidelines for the description of mainland Southeast Asian languages. In A. Vittrant & J. Watkins (Eds.), *The Mainland Southeast Asia Linguistic Area* (pp. 653–686).

von Prince, K. (2017). Dozing eyes and drunken faces: Nominalized psycho-collocations in Daakaka (Vanuatu). *Studies in Language, 41*(2), 502–520. https://doi.org/10.1075/sl.41.2.08von

Weißmann, L. (2019). *Hard hearts, soft hearts, swollen hearts. Experiental collocations with ini ('heart') in Dzaha Dzaui (colonial Mixtec) in Francisco de Alvarado's "Vocabulario en lengua misteca" (1593)* [M.A., Freie Universität Berlin]. https://doi.org/10.13140/RG.2.2.28319.10408

Weissmann, L. (2023). Psycho-Collocations with Ini ‘Heart’ in Teposcolula Mixtec. In J. Baranyiné Kóczy & S. Katalin (Eds.), *Embodiment in Cross-Linguistic Studies* (pp. 415–438). Brill. https://doi.org/10.1163/9789004548640_019

Westermann, D. (1973). *Evefiala or Ewe–English dictionary: Gbesela yeye or English–Ewe dictionary*. Kraus Reprint.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer. https://ggplot2.tidyverse.org/

Wierzbicka, A. (1999). *Emotions across languages and cultures: Diversity and universals*. Cambridge University Press.

Wilkins, D. P. (1996). Semantic change and the search for cognates. *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*, 264–304.

Yandt, G. (2019). Variation and culture in Laiholh psycho-collocations. *Indiana Working Papers in South Asian Languages and Cultures, 1*(1), 18.

Ye, Z. (2002). Different modes of describing emotions in Chinese: Bodily changes, sensations, and bodily images. *Pragmatics & Cognition, 10*(1–2), 307–339.

Zoch, U. (2020). Experiencer constructions. In R. Vossen & G. J. Dimmendaal (Eds.), *The Oxford handbook of African languages*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199609895.013.54

Appendices

Appendix A. Spearman's rank correlation between frequency of person-part and number of subdomains in entries with that person-part for each language

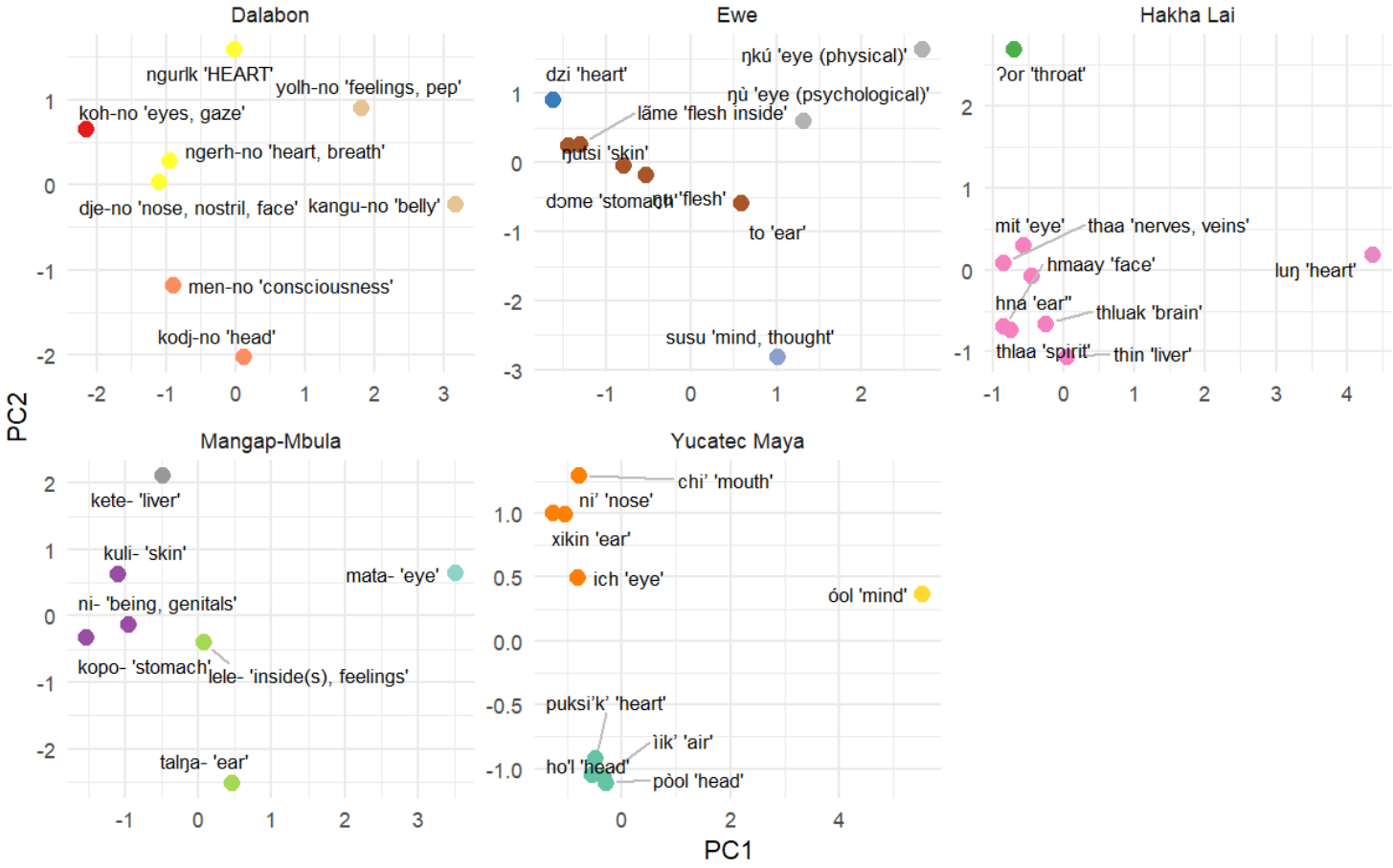
The following table represents the Spearman's rank correlation between frequency of person-part and number of subdomains in entries with that person-part for each language. Note p-values in the table are based on large-sample approximation given tied ranks. Further, note that this analysis is largely unchanged when excluding all database entries coded for more than one subdomain. The only notable difference between this analysis and the results given in the table below are that the correlation found for Dalabon is no longer statistically significant (Spearman's r = 0.66, df = 6, p=.104), which can be explained by a skew toward the emotion subdomain for Dalabon entries in the database (cf. §3.2).

Language	***Correlation coefficient (Spearman's r)***	***Degrees of freedom (df)***	***P-value (p)***
Dalabon	0.87	6	.005
Hakha Lai	-0.19	7	.630
Ewe	0.37	5	.323
Mangap-Mbula	0.83	5	.021
Yucatec Maya	0.71	7	.032

Appendix B. Clustering of person-parts by subdomain in each language

The figure below represents principal components analysis (PCA) derived from assigned experiential subdomain(s) for each person-parts within each language in the database. Points are coloured according to hierarchical clustering on principal components (HCPC). These figures were generated in R using the FactoMineR and factoextra packages for PCA and HCPC (Lê et al 2008; Kassambara and Mundt 2020) as well as the ggplot2, ggrepel and RColorBrewer for graphing (Neuwirth, 2022; Slowikowski, 2024; Wickham, 2016).

When viewing the below, there are a few things to keep in mind. Clusters within a given language are independent of clusters given for any other language. The distance between points represents similarity following the two principal components plotted on the x-axis and y-axis. Finally, note that HCPC (clustering represented by colours of points) may be derived from more principal components than are represented by the distance between points below.


Dalabon
ngurlk 'HEART'
yolh-no 'feelings, pep'
koh-no 'eyes, gaze'
ngerh-no 'heart, breath'
dje-no 'nose, nostril, face'
kangu-no 'belly'
men-no 'consciousness'
kodj-no 'head'
Ewe
ŋkú 'eye (physical)'
dzi 'heart'
ŋù 'eye (psychological)'
lãme 'flesh inside'
ŋutsi 'skin'
dɔme 'stomach'
nu 'flesh'
to 'ear'
susu 'mind, thought'
Hakha Lai
ʔor 'throat'
mit 'eye'
thaa 'nerves, veins'
hmaay 'face'
luŋ 'heart'
hna 'ear'
thluak 'brain'
thlaa 'spirit'
thin 'liver'
Mangap-Mbula
kete- 'liver'
kuli- 'skin'
mata- 'eye'
ni- 'being, genitals'
kopo- 'stomach'
lele- 'inside(s), feelings'
talŋa- 'ear'
Yucatec Maya
chi' 'mouth'
ni' 'nose'
xikin 'ear'
ich 'eye'
óol 'mind'
puksi'k' 'heart'
ìik' 'air'
ho'l 'head'
pòol 'head'
PC2
PC1

Appendix C. Contingency table between subdomain and clusters given in Figure 5.4

The below table shows combines database entries with respect to the three clusters of person-parts given in Figure 5.4. These clusters are then distributed with respect to the experiential subdomains associated with entries in each cluster. Note that clusters were determined based on the experiential subdomains associated with unique pairings between psycho-mate and psycho-noun (cf. §5.2.1).

		Experiential Subdomain				
		Cognition	*Perception*	*Volition*	*Emotion*	*Bodily Sensation*
Cluster	*1*	75	18	70	270	134
	2	28	0	1	2	1
	3	5	39	9	10	4