LUND UNIVERSITY

Recurrent Gestures, Embodiment and Sedimentation

The Case of Two-Handed-Alternation-Sagittal (2HAS) Gesture in Japanese and Swedish

Daiya Kuryu

Supervisor: Jordan Zlatev

Centre for Language and Literature, Lund University

MA in Language and Linguistics, Cognitive Semiotics

SPVR01 Language and Linguistics: Degree Project – Master's (Two Years) Thesis, 30 credits

May 2025

Abstract

This thesis investigates the semiotic foundations of recurrent gestures through a cross-cultural study of the Two-Handed-Alternation-Sagittal (2HAS) gesture in Japanese and Swedish. Drawing on the Motivation & Sedimentation Model (MSM)—a cognitive-semiotic model of meaning-making—the study explores how recurrent gestures emerge from the dynamic interplay between embodiment and sedimentation. Accordingly, the thesis describes an empirical study aiming to what extent recurrent gestures, situated between spontaneous gesticulation and conventionalized emblems, are shaped both by shared bodily experiences and by historically sedimented social practices. The 2HAS gesture, characterized by alternating hand movements along the sagittal axis, was analyzed for its meaning variations and distribution using phenomenological triangulation: combining first-person conceptual reflection, second-person intersubjective methods, and third-person quantitative analysis. Through an examination of semi-spontaneous conversations in two groups of participants, Japanese and Swedish, the study reveals both commonalities and differences in the use and meaning of 2HAS across these cultural groups. The gesture frequently expressed delineated meanings like BIDIRECTIONALITY and GIVE & RECEIVE, as well as holistic meanings like UNCERTAINTY, OPTIONALITY and WORD SEARCH. While commonalities support the gesture's grounding in pan-human aspects of embodiment, cultural variations point to the influence of culture-specific sedimentation. These findings refine our understanding of the nature of recurrent gestures and offer empirical grounding for a revised version of MSM, which incorporates a hybrid Embodied/Sedimented level and further distinguishes genetic (individual) and generative (historical-social) layers of sedimentation.

Keywords: cognitive semiotics, recurrent gestures, embodiment, sedimentation, Two-Handed-Alternation-Sagittal (2HAS) gesture, the Motivation & Sedimentation Model (MSM)

Acknowledgements

This thesis would not have been possible without the support of countless people, only some of whom I can mention here due to space limitation.

First and foremost, I would like to express my deepest gratitude to my supervisor, Jordan Zlatev, for his unwavering support throughout my academic journey at Lund. When I first read his paper on *mimetic schemas* (Zlatev, 2005) and his introductory article on cognitive semiotics (Zlatev, 2012) more than three years ago, I was just a student aspiring to become a cognitive linguist with an interest in cognitive semiotics. Now, at the end of this two-year MA program, thanks to his guidance, I consider myself a junior cognitive semiotician with a background in cognitive linguistics. Although my academic career is still in its early stages, I am confident that the solid foundation I have built in cognitive semiotics will continue to guide and inspire me throughout my future endeavors. Equally heartfelt thanks go to my Swedish research partner, Harry Polfeldt, without whom this thesis would *literally* not have been possible. I benefited immensely from his expertise in gesture studies and found myself growing intellectually through our collaboration.

I would also like to express my appreciation to my fellow MA students in linguistics and the PhD students in cognitive semiotics. In particular, I want to thank Jahnesta Sjöberg and Saga Bringeland for the stimulating discussions that made my time in Lund both academically enriching and enjoyable. My thanks also extend to those I had the pleasure of meeting at conferences and through emails. In particular, I would like to acknowledge Silva Ladewig for her insightful feedback on the 2HAS recurrent gesture. I am also grateful to all the participants who took part in this study—beyond the 2HAS gesture, there are many other fascinating gestures that continue to inspire ideas for future research.

Lastly, to my family and my partner, Haoyang: from chatting with me to helping me navigate life decisions, your support has been indispensable—past, present, and future. I promise to repay all you have given me by pursuing my dreams, the next of which is to obtain a PhD on the fascinating phenomenon of gesture, spoken language, and their interplay—*polysemiosis*.

Table of contents

List of figures

List of tables

Abbreviations

2HAS	Two-Handed Alternation gesture on the Sagittal axis
CI	Communicative Intent

CrI	Credible Interval
MSM	Motivation & Sedimentation Model
PUOH	Palm Up Open Hand
RC	Representational Complexity
RQ	Research Question
SD	Standard Deviation

Chapter 1. Introduction

When we communicate with others, we do so not only by language but also through our bodies—movements of the speaker's hands, arm, torso, head, and whole body with "features of manifest deliberate expressiveness"—those that are subsumed under the term "gestures" (Kendon, 2004, pp. 7, 13-14; cf. Müller, 2014). Some gestures, such as THUMBS UP and CROSSED FINGERS, are highly conventionalized and adhere to established standards of form, carrying word-like meanings. These are referred to as *emblems* or "quotable gestures" (Kendon, 2004). On the other hand, as is often assumed, other gestures seem to be free from such standards of form, created "on the fly" depending on the local context. Some scholars who take a psychological approach have almost exclusively focused on such "spontaneous gestures", arguing that they provide windows onto mind (McNeill, 1992).

However, apart from emblems, are all gestures spontaneous in the way envisioned by McNeill (1992)? Are there not any regularities in the way we gesture? The last two decades have seen an upsurge in the line of research that explores these questions. For instance, when discussing ongoing events or the repetition of events, speakers often produce gestures with cyclic movements alongside their speech. These "cyclic gestures" are also systematically used when the speaker is searching for a word or concept in their mind (Ladewig, 2011). These findings support the idea that certain gestural forms are associated with specific, yet not too narrow, semantic and pragmatic meanings. However, unlike emblems, these gestures are not fixed in form; rather, they exhibit formal variations depending on their specific meanings in local contexts. These gestures are referred to as *recurrent gestures* and are often defined as "*partly* conventionalized gestures" (Ladewig, 2025, p. 2, my emphasis) for the very reasons outlined above.

Recurrent gestures are thus repeatedly used across speakers "within a given speech-gesture community" (Müller, 2017, p. 280). This could imply that different kinds of recurrent gestures are to be found in different cultures, as have been highlighted by anthropological studies: "gestures reveal a great deal about interactional practices, the social norms that underlie them and how local and wider ideologies in societies shape the nature of gestures and their use" (Brookes & Le Guen, 2019, p. 129). On the other hand, the recurrency could be due to more general, pan-human embodied experiences, such as basic manual actions or visual perception of and bodily experiences with movements (Müller, 2017). Given that these embodied experiences are more or less universal, it stands to reason that there would also be some commonalities in the uses of recurrent gestures across different speech communities.

Thus, it remains an open question whether recurrent gestures are primary grounded in conventionality, and thus the semiotic ground of *symbolicity*, or rather that of *iconicity*, or perhaps a combination of both; correspondingly, though not synonymously, the question can be posed whether recurrent gestures are to a greater extent motivated by cultural *sedimentation*, or by pan-human *embodiment* (see Chapter 2 for explanations of these concepts). The present thesis seeks to explore these issues, by comparing the use of a specific recurrent gesture between Japanese and Swedish people.

The basic methodological premise of the present study is that to the extent that a recurrent gesture is motivated by sedimented social beliefs and norms, it will exhibit culture-specific aspects. In the same vein, the more it is grounded in our pan-human embodied experiences, the more commonalties it will exhibit across different cultural groups. Furthermore, from a theoretical viewpoint, pursuing a cross-linguistic study in this direction will necessitate a re-conceptualization of the very notion of recurrent gestures, as it concerns a fundamental question that has hitherto been neglected in the previous literature: Where does "recurrency" come from? Is it because of social sedimentation, because of pan-human embodiment, or perhaps both?

To investigate this principal research question, this thesis employs methods and concepts from the discipline of *cognitive semiotics* (Zlatev, 2012, in press). With insights from phenomenology (Sokolowski, 2000; Zahavi, 2003, 2018), cognitive semiotics places primary emphasis on subjectivity—"any kind of qualitative experience" that is irreducible to any physicalist ontology (Zlatev, in press, p. 4). This aligns well with the present study, as the principal research question concerns the *essence* of recurrency, which necessitates the researcher's qualitative analysis of the phenomenon, such as *eidetic variation* (Sokolowski, 2000; Zahavi, 2018). In addition to such first-person methods, cognitive semiotic research also employs second-person and third-person methods: *phenomenological triangulation* (see Section 2.1) This study applies such triangulation by incorporating both an intersubjective corroboration step between researchers and a more detached quantitative analysis (see Chapter 3 for further methodological details). Furthermore, it provides an analysis based on the Motivation & Sedimentation Model (MSM), a recent cognitive-semiotic theory that integrates both culture-specific and pan-human aspects of meaning-making (see Section 2.1.4 for more details). With these key notions from cognitive semiotics, this thesis aims to return at the end to the fundamental concepts of recurrent gestures, sedimentation, and embodiment, incorporating the empirical findings of the present study. This continuous refinement of

concepts through empirical findings is known as the *conceptual-empirical loop*, a core principle in cognitive semiotics (Zlatev, 2015a).

With this background, this thesis examines the central research question by comparing semi-spontaneous dialogues made by Japanese and Swedish participants, with a particular focus on a specific recurrent gesture: Two-handed Alternation gestures on the Sagittal axis (2HAS; Kuryu, 2024a, 2024b, 2025). Despite its inclusion in a list of recurrent gestures (Bressem & Müller, 2014b, p. 1580), 2HAS remains understudied in the literature. Building on my previous studies on conventionalized polysemiotic units consisting of linguistic and gestural components (known in the literature as "multimodal constructions", Zima & Bergs, 2017), this thesis also aims to provide a comprehensive analysis of the meaning of 2HAS as a recurrent gesture. The investigation is thus guided by the following more specific research questions:

RQ1: What meaning variations does 2HAS exhibit, particularly from a cross-cultural perspective?

RQ2: To what extent is 2HAS motivated by embodiment and to what by sedimentation, as reflected in its commonalities and differences across the cultural groups, respectively?

The remainder of this thesis is structured as follows: Chapter 2 presents the theoretical background, offering a more detailed description of recurrent gestures and MSM within the broader contexts of gesture studies and cognitive semiotics, respectively. The chapter also reviews my previous studies concerning 2HAS (Kuryu, 2024a, 2024b, 2025). Chapter 3 elaborates on the methods of this empirical study in light of phenomenological triangulation. Chapter 4 presents the results and discussion of the cross-cultural analysis of 2HAS, and Chapter 5 concludes the thesis by summarizing the findings and their theoretical implications.

Chapter 2. Theoretical background

Chapter 2 introduces the theoretical framework underpinning this thesis. Section 2.1 first presents key concepts from cognitive semiotics—particularly embodiment, sedimentation, and the Motivation & Sedimentation Model (MSM)—and explains how these relate to the study of gestures. Section 2.2 then elaborates on gesture classifications from both semiotic and functional perspectives, with a focus on recurrent gestures. Finally, Section 2.3 reviews previous studies, including my own research, on the 2HAS gesture. Together, these discussions provide a conceptual foundation for examining how 2HAS gestures are motivated by pan-human embodiment and/or shaped by culture-specific sedimentation.

2.1. Cognitive semiotics

Cognitive semiotics has emerged over the past two decades as a new transdisciplinary science, shaped through collaborations among linguists, semioticians, and cognitive scientists, and its definition has also developed accordingly (e.g., Brandt, 2011; Sonesson 2012; Zlatev, 2012, 2015). Most recently, Jordan Zlatev, the first president of the *International Association for Cognitive Semiotics* (https://iacs.dk), defined it as "the academic discipline that focuses on meaning-making (semiosis), combining concepts and methods from semiotics, cognitive science, linguistics, as well as phenomenology" (Zlatev, in press, p. 2). In addition to the three main contributing disciplines, Zlatev (in press) notably includes the philosophical tradition of phenomenology, which is "the study of human experience and of the ways things present themselves to us in and through such experience" (Sokolowski, 2000, p. 2). This inclusion is not merely additive; rather, phenomenology plays a central role in synthesizing the other three disciplines into a unified study of meaning and meaning-making—without falling into reductive dichotomies such as "body vs. mind" or "mind vs. culture". This influence from phenomenology is reflected in the methodological principles of cognitive semiotics, which place primary emphasis on (inter)subjectivity.

One of the methodological principles practiced in cognitive semiotics is *phenomenological triangulation* (e.g., Mouratidou, in press; Pielli & Zlatev, 2020; Zlatev, 2012, 2015, in press). Studies employing this triangulation integrate methods from the first-person, second-person and, if necessary, third-person perspectives, following this specific order of precedence (and thus methodological focus) (Mendoza-Collazos & Zlatev, 2022; Zlatev & Mouratidou, 2024). This explicit prioritization of first- and second-person methods—that is,

our lived experiences as beings-in-the-world—is succinctly expressed by Gallagher and Zahavi (2008, pp. 90–91) as follows:

> We should never forget that our knowledge of the world, including our scientific knowledge, arises from a first-person perspective, and that science would be meaningless without this experiential dimension. […] Scientific knowledge depends (although, of course, not exclusively) on the observations and experiences of individuals: it is knowledge that is shared by a community of experiencing subjects.

Methods from a first-person perspective involve the researcher's conceptual analysis of the target phenomenon, grounded in their own intuitions (Zlatev & Blomberg, 2019), which are refined through prior experiences with the phenomenon. These first-person methods are followed by second-person methods, which emphasize the interpersonal dimension of the investigation, often involving collaboration with co-researchers or direct interaction with participants. Finally, third-person methods adopt the most detached and observational stance, where the researcher aims to measure, observe, or analyze the phenomenon—primarily in a quantifiable manner, such as through psycholinguistic experiments and statistical analysis. Table 1 summarizes the perspectives and associated methods of phenomenological triangulation, and Section 3.1 outlines how this triangulation is applied in the present study.

Table 1. Perspectives and related methods of phenomenological triangulation (based on Zlatev, 2012)

Perspective	**Methods**
First-person ("subjective")	Conceptual analysis Phenomenological methods Systematic intuitions
Second-person ("intersubjective")	Empathy Imaginative projection
Third-person ("objective")	Detached observation Experimentation Computational modelling

Another methodological principle in cognitive semiotics is the *conceptual-empirical loop* (e.g., Devylder & Zlatev, 2020; Mendoza-Collazos & Zlatev, 2022; Stampoulidis et al., 2019;

Zlatev, 2015a). As highlighted in the previous paragraph, cognitive semiotic research typically begins with a first-person, conceptual analysis of the target phenomenon. However, unlike traditional semiotics—which "has often proposed rather general, and empirically under-supported theories" (Mendoza-Collazos & Zlatev, 2022, p. 143)—cognitive semiotics incorporates empirical methods, guided by prior conceptual considerations. The resulting empirical findings are then fed back into the conceptual analysis, allowing for a refinement and deepening of our understanding of the phenomenon—thus forming a continuous loop. For the present thesis, the main concepts to which the conceptual-empirical loop is applied are *recurrent gesture*, *embodiment*, and *sedimentation*. The conceptual side of the loop is developed in the subsequent subsections of this chapter, through a thorough review of previous literature on these concepts. Based on this foundation, an empirical study focusing on the specific recurrent gesture of 2HAS is presented in Chapters 3 and 4. In the latter part of Chapter 4, we return to the conceptual discussion, revisiting and refining the key concepts in light of the empirical findings. Thus, this thesis is structured according to the conceptual-empirical loop.

2.1.1. Consciousness, semiosis and sign

Moving from the key methodological concepts in cognitive semiotics, this subsection now turns to the central theoretical concepts relevant to the present thesis.

Given the influence of phenomenology on cognitive semiotics, as outlined in the previous section, it is natural to begin with the notion of *intentionality*, understood as the directedness of consciousness toward an object, or even more generally, beyond itself. Consciousness, in this view, is characterized by its inherent "pointing-beyond-itself" nature (Zahavi, 2018, p. 16)—our "openness to the world", in the words of Merleau-Ponty. Another key concept, *semiosis*, is not merely the use or manipulation of signs, as often conceived in traditional semiotics, but is more broadly understood as meaning-making. In light of phenomenological insights, particularly from Merleau-Ponty, consciousness and semiosis are not reducible to one another, but are rather co-constituted dimensions of what it means to be a subject in the world. As Merleau-Ponty famously writes:

> The world is inseparable from the subject, but from a subject which is nothing but a project of the world, and the subject is inseparable from the world, but from a world which the subject itself projects. The subject is a being-in-the-world [être au monde]. (Merleau-Ponty 1962: 499–500)

Figure 1 illustrates this reciprocal, albeit asymmetric, subject-world relationship, where "consciousness (intentionality) [is] the more world-directed aspect of existence, while semiosis [is] the more subject-directed ('inward') side" (Zlatev & Konderak, 2022, p. 170).

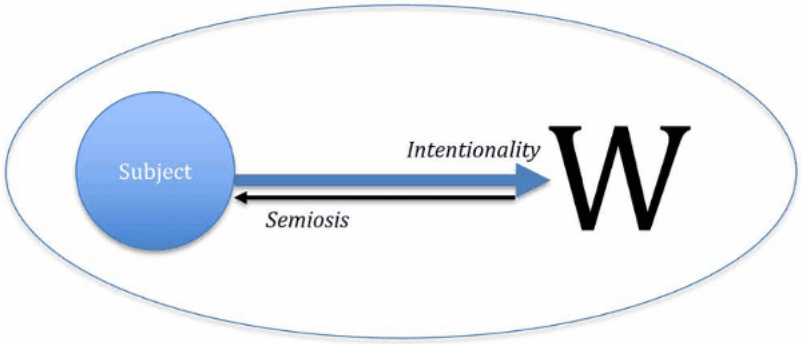


Figure 1. The reciprocal relations between intentionality and semiosis, with W indicating the life world (reprinted from Zlatev & Konderak, 2022, p. 171)

Building on this dynamic interrelation, a recent version of the *Semiotic Hierarchy* framework (Zlatev & Konderak, 2022)—a stratified model of meaning-making—was developed from earlier versions (Zlatev, 2009, 2018). It comprises five interlinked layers, each characterized by a specific form of intentionality and corresponding structures of semiosis, as shown in Table 2.

Table 2. A recent version of the Semiotic Hierarchy (reprinted from Zlatev & Konderak, 2022, p. 178)

Layer and label	**Intentionality**	**Structures of semiosis**
5. Language	Linguistic intentionality	Predication, syntax
4. Sign use	Signitive intentionality	Signs
3. Intersubjectivity	Shared intentionality	Mimetic schemas, conventions, rituals
2. Subjectivity	Perceptual intentionality	Noemata, Leib/Körper duality
1. Animation	Operative intentionality, drive intentionality	Schemas, goals, emotions

The first layer is grounded in *operative intentionality*, which "includes a constant 'affective

tone' accompanying all acts of movement and perception", serving as "a non-representational *background of subjectivity*" (Zlatev & Konderak, 2022, p. 179, emphasis original). In other words, there is "a level of experience that is fundamentally tied to the living body" (Zlatev, 2018, p. 7), which works at the margins of consciousness. The second layer introduces *perceptual intentionality* and pre-reflective self-consciousness (Zlatev, 2018). The consciousness-semiosis unity at this level is analyzed in terms of the *noesis–noema* structure, referring to the intentional act and its intentional object, respectively. The third layer is characterized by *shared intentionality* and empathy. This enables the co-construction of meaning: communication. However, Zlatev (2018) stresses that while this level enables social meanings, it does not yet require the use of signs: "there are social meanings such as joint actions, types/categories, and even conventions that are not signs, strictly speaking" (p. 2). The fourth layer is defined by (non-linguistic) *signitive intentionality*, where expression and content are differentiated by a conscious interpreter. This relies on what Sonesson calls "double asymmetry": the expression is more directly experienced, while the content is more thematically central (Sonesson, 2010). The highest layer is marked by *linguistic intentionality*, which distinguishes it from other types of sign use in that linguistic symbols "are 'designed' to combine in articulated structures, with the help of 'syncategorematic' (i.e. grammatical) elements to form linguistic constructions such as predication, modification, conjunction and subordination" (Zlatev & Konderak, 2022, p. 185).

The vertical relationship among these layers is best understood through the phenomenological notion of *Fundierung*, "with lower levels prefiguring the higher ones and higher ones consolidating and sublimating the lowers ones, but without breaking away from them" (Zlatev, 2018, p. 6). As Merleau-Ponty describes, "the originated [i.e. the higher] is presented as a determinate or explicit form of the originator [i.e. the lower], […] and yet it is through the originated that the originator is made manifest" (1962, p. 458). Meaning-making is also shaped by a dynamic tension between sedimentation and spontaneity: norms and structures arise from spontaneous creative acts, but in turn constrain future acts without fully determining them.

With the help of the Semiotic Hierarchy framework, we can clearly see that *signs* are only one type of structures of semiosis. Still, the concept of the sign remains central in cognitive semiotics and involves, according to Sonesson, the key properties of differentiation of expression and content: "double asymmetry between the two parts, because one part, expression, is more directly experienced than the other; and because the other part, content, is

more in focus than the other" (2010, p. 25).

In agreement with this, but in a more compact manner, Zlatev et al. (2020, p. 160, italics removed) propose a definition of sign use grounded in reflective, signitive intentionality, as a consequence of which the "double asymmetry" emerges:

> DEF. A sign <E, O> is used (produced or understood) by a subject S, if and only if:
> (a) S is made aware of an intentional object O by means of expression E, which can be perceived by the senses.
> (b) S is (or at least can be) aware of (a).

This definition implies that the sign use requires not only the subject's perceptual consciousness of the *expression* E and its relation with an *intentional object* O (which can be anything that consciousness is directed to, concrete or abstract, existent or inexistent, etc.), but also their potential to be reflectively aware of that very relation between E and O. As Zlatev and Konderak (2022, p. 183) note, it is precisely this reflective awareness "that distinguishes associative signals [e.g., vervet monkeys' alarm calls] from signs proper and provides the 'double asymmetry' in question as a by-product."

Regarding the *kind* of relation between E and O, Sonesson's (2007, 2010) interpretation of Peirce's notion of *ground* (Peirce, 1960: CP 2.228) is of particular relevance. According to Sonesson (2007), the term "ground" refers to "those properties of the two things entering into the sign function by means of which they get connected, i.e. both some properties of the thing serving as expression and some properties of the thing serving as content" (p. 24). In short, it concerns the relation between E and O, which Sonesson (2007, 2010), following Peirce, classifies into three types: *iconic* (based on the similarity between E and O), *indexical* (based on contiguity in space or time between E and O), and *symbolic* (based on convention, i.e., a shared agreement in a community that E signifies O). These three *semiotic grounds* are not mutually exclusive and thus can coexist within a single sign, although typically one tends to predominate over the others, as noted by Jakobson (1965). For example, both THUMBS UP 👍 and THUMBS DOWN 👎 gestures are regarded as conventionalized gestures, or so-called emblems, as introduced in Chapter 1 (see also Section 2.2). The former signifies POSITIVITY and the latter NEGATIVITY, and these are mainly due to social conventions, i.e., the symbolic ground predominates. However, these E-O relations may also involve iconic motivations—particularly *diagrammatic iconicity* (Peirce, 1903: CP 2.277)—as they are rooted in our bodily

movement along the vertical axis, through which we experientially correlate affective valence: standing upright when healthy or victorious versus slumping when defeated, sad, or sick (Lakoff & Johnson, 1980).[1]

In addition to the semiotic ground, another important dimension of the E-O relation concerns how an expression approaches or *construes* an intentional object (Möttönen, 2016; Zlatev, 2016; Zlatev & Möttönen, 2023). In other words, an intentional object does not present itself in a neutral manner; rather, it is always given to a conscious subject in a particular, perspectivized way, depending on the expression. For a linguistic example, Zlatev (2016) draws on Frege's (1948) well-known case of the *Morning Star* and the *Evening Star*, where the same intentional object is construed differently. Section 2.3.2 will demonstrate how this applies to gestural expressions.

In sum, through *signitive intentionality*, the conscious subject perceives the expression and understands its relation to the signified intentional object, with the link based on the semiotic ground of iconicity, indexicality, symbolicity, or a combination thereof. As a form of semiosis directed back to the subject, the intentional object manifests in a construal shaped by the expression. Figure 2 illustrates the comprehensive concept of the sign, encompassing all the main concepts introduced in this section.

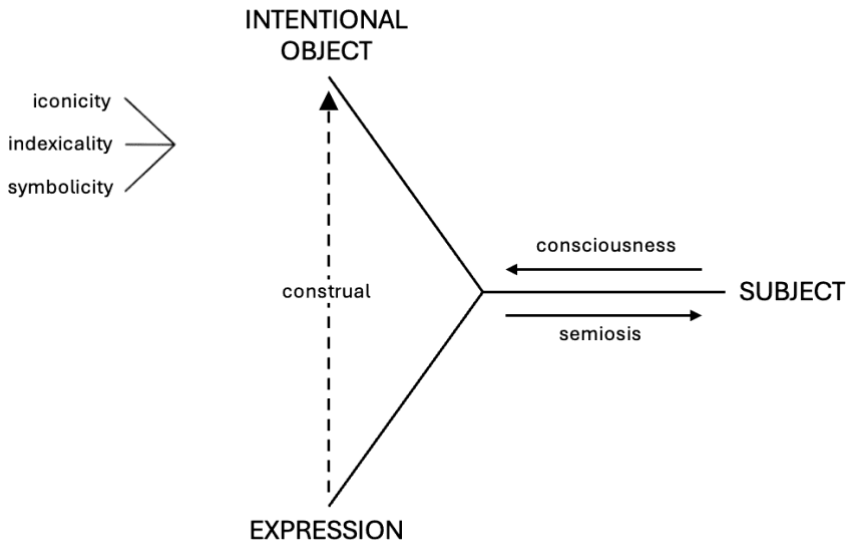


Figure 2. A representation of the various elements involved in sign use (adapted from Højgaard Hansen 2024, Figure 5)

[1] While these motivations are understood as "conceptual metaphors" in the field of cognitive linguistics (Lakoff & Johnson, 1980, 1999), these are not considered metaphors from a cognitive-semiotic perspective, as they are not signs (Zlatev, 2024).

2.1.2. Embodiment

In the preceding section I briefly outlined the Semiotic Hierarchy framework (Zlatev, 2009, 2018; Zlatev & Konderak, 2022), which distinguishes five inter-dependent layers of meaning—animation, subjectivity, intersubjectivity, sign use, and language—bound together by *Fundierung*, a relation in which each higher layer are "sublimated" from but never reducible to the lower one (Zlatev, 2018). In what follows I clarify how *phenomenological embodiment*—understood after Merleau-Ponty as the lived, sense-making body—constitutes the experiential core of every layer. Whereas Rohrer (2007) and Pielli and Zlatev (2020, pp. 6-10) survey multiple senses of the term "embodiment", I focus on the phenomenological strand that privileges first-person bodily experience over neural or "cognitive-unconscious" (Lakoff & Johnson, 1999) approaches, in line with the primacy of subjectivity in cognitive semiotics.

At the basal animation layer, meaning is enacted by operative intentionality—the pre-reflective grip the living body has on its surroundings. This grip is realized by the *body schema*, "a system of sensory-motor processes that constantly regulate posture and movement […] that function without reflective awareness or the necessity of perceptual monitoring" (Gallagher 2005 pp.37-38). In other words, through this body schema, "the body 'adapts' to the invitation of the world" (Moya, 2014, p. 2).

Ascending to the next layer, this layer affords a richer articulation of the *Leib/Körper* duality, a central concept in phenomenology. While *Körper* refers to the objective, material body, *Leib* denotes the lived, subjective body—the body as it is experienced from within. This duality becomes especially salient in the phenomenon of "double sensation," vividly described by Merleau-Ponty (1962, p. 106):

> [W]hen I touch my right hand with my left, my right hand, as an object, has the strange property of being able to feel too. […] When I press my two hands together, it is not a matter of two sensations felt together as one perceives two objects placed side by side, but of an ambiguous set-up in which both hands can alternate the rôles of 'touching' and being 'touched.'

Importantly, this duality of the body serves as an essential precondition for the next layer of intersubjectivity through empathy, as described by Zahavi (2003, p. 104):

> Thus, it is exactly the unique subject-object status of the body, the remarkable interplay between ipseity and alterity characterizing double-sensation, which permits me to recognize and experience other embodied subjects. […] I am experiencing myself in a manner that anticipates both the way in which an Other would experience me and the way in which I would experience an Other. […] The possibility of sociality presupposes a certain intersubjectivity of the body.

At this third layer perceptual intentionality is now sublimated to shared intentionality, forming a foundation for *bodily mimesis* (Donald, 1991; Zlatev, 2007, 2008), "the volitional use of the body to first imitate the movements and actions of *others*, and eventually […] to re-enact particular actions and events (and possibly even narratives) to *others*" (Zlatev, 2018, p. 11, my emphasis). On the basis of this bodily mimesis, (non-present) actions and events become internalized as *mimetic schemas*—body-based, pre-linguistic experiential structures that are consciously accessible and pre-reflectively shared in a community (Zlatev, 2005, 2014). Crucially, mimetic schemas can be enacted both covertly (in imagination) and overtly (as iconic gestures) (Zlatev, 2014), thereby serving as a scaffold for the next two layers: sign use and language.

In connection with mimetic schemas, it is worth mentioning another type of embodied schema that has been widely utilized in cognitive-linguistic studies: the *image schema*. Johnson (1987) defines an image schema as "a recurring, dynamic pattern of *our* perceptual interactions and motor programs that gives coherence and structure to *our* experience" (p. xiv, my emphasis). Unlike mimetic schemas, however, the main proponents of image schemas have argued that these structures operate beneath the level of conscious awareness—that is, within what Lakoff and Johnson (1999) term the "cognitive unconscious." Yet, this cognitive-unconscious account appears incompatible with their allegedly "public, shared meaning" (Johnson, 1987, p. 190), as also implied by the use of the first-person plural possessive pronoun in Johnson's definition. Drawing on findings that mimetic schemas ontogenetically precede image schemas (Zlatev, 2014), this thesis adopts the perspective that "image schemas come about as generalizations from the more concrete structures of embodied intersubjectivity [such as mimetic schemas], and may thus inherit their interpersonal aspects this way" (Zlatev, 2017, p. 180). Importantly, this perspective entails that image schemas are not operative in the cognitive unconscious, but at least at the margins of consciousness.[2]

[2] Alternatively, drawing on Langacker's (2006) interpretation, Zlatev (2010) offers another

As implied by the interpersonal nature of these embodied schemas, it is important to emphasize that embodiment and intersubjectivity are complementary phenomena, as highlighted by Zahavi's (2003, p. 104) quote above. Accordingly, the meaningfulness of embodied structures and processes discussed so far is "not private, but interpersonal—not 'in the head' but 'in the world'" (Zlatev, 2017, p. 176). This *embodied intersubjectivity* (or intercorporeality) serves as a foundational substrate upon which the highest two layers—sign use and language—are sedimented (Zlatev & Blomberg, 2016). The concept of *sedimentation* is further explored in the following section.

Finally, embodiment also explains the cross-cultural commonality of basic meaning-making. As Pelkey (2023, p. 8) aptly puts it:

> Presumably, then, since human beings around the world share a common embodiment (upright posture, bipedal loco-motion, lateralized specialization, descended larynx, opposing thumbs, fine manual motor control, and so forth—in addition to our primate, mammalian, and vertebrate layers of shared existence), some aspects of our lived experience in the world may reasonably be expected to transcend cultural and regional conditioning, enabling a shared substrate for mutual understanding, learning, translating, and communicating.

These shared bodily invariants ensure that the pan-human embodied processes and structures reviewed in this section[3] provide the basis for shared human meaning-making—a point particularly relevant to gestures, a semiotic system whose semiotic ground is predominantly iconic.

2.1.3. Sedimentation

In the previous section, I discussed how pre-linguistic and pre-signitive meaning, located at the lower levels of the Semiotic Hierarchy framework, emerges through (phenomenological) embodiment. In this chapter, I shift the focus to the highest two levels of the Semiotic

phenomenological take on image schemas: "Langacker is proposing to regard 'image schemas' not as structures, representations etc., but as *processes* of (human) consciousness through which experience is 'analyzed', and which therefore […] are in the background of consciousness" (p. 18, emphasis original). This view suggests that, continuous with body schemas, image schemas function at the margins of consciousness as a form of operative intentionality.

[3] Regarding mimetic schemas, while some—such as GIVE and RECEIVE (see also Section 2.3.2)—are considered more or less universal, others, like EAT and SIT, exhibit culture-specific aspects.

Hierarchy—namely, signs and language—and examine how these, initially arising as creative acts (spontaneity) grounded in embodied motivation, are transformed into socially shared structures—a process known as sedimentation.

The concept of *sedimentation* originates in Husserl's phenomenology and describes how meanings progressively become abstracted, stabilized, and conventionalized through symbolic systems such as language and mathematics (Blomberg, 2019). In its most basic sense, sedimentation refers to the passive retention of past experiences, which later inform and shape new ones. As Sonesson (2021) explains, "such sedimentation may be *genetic* (derived from the experience of the individual from the cradle to the grave – or almost) or *generative* (handed down from one generation to another, and so on indefinitely)" (Sonesson, 2021, p. 117, my emphasis).

Genetic sedimentation emerges from the habitualization of personal embodied interactions and experiences, forming stable structures such as habits and schemas within an individual's cognitive and bodily repertoires. It involves the ongoing, tacit accumulation of past experiences, providing "a cognitive scaffold that liberates the thinking of the human individual from the impossible task of thinking everything simultaneously and constantly anew" (Woelert, 2011, p. 120). Generative sedimentation, rather, operates on a broader scale, encompassing historical and cultural dimensions. It results from the cumulative and collective transmission of meaning across generations, embedding conventions into socially accessible structures. The metaphor of sedimentation aptly captures this vertical accumulation over historical time, resulting in stable horizontal social strata, as summarized by Woelert (2011, p. 119):

> [T]he metaphor of sedimentation allows one to imaginatively connect the two dimensions of conceptual-cognitive processes that are often separated: their synchronic structure on the one hand, their diachronic dynamics on the other. Sedimentation, in spatial terms, describes a process whereby particles collect together and build vertically. This vertical process, in turn, leads to the establishment of horizontal strata that over time form a stable structural configuration.

(Generative) sedimentation is two-faced. On the one hand it enables stability of meaning "across space and time"; on the other it "entails a forgetfulness of the origin that made the sense in question possible in the first place" (Blomberg, 2019, p. 79). The semiotic ground of symbolicity—introduced in Section 2.1.1—thrives precisely on this abstraction: once a sign's

bodily motivations fade, its normative, symbolic value can dominate. Precisely because a symbol works by convention, we can use it without re-enacting the embodied motivations that once animated it (Blomberg & Zlatev, 2014; Blomberg, 2019).

Consider the emblematic THUMBS UP gesture, first mentioned in Section 2.1.1. The pioneer who first raised a thumb to signify APPROVAL may have relied on several bodily motivations: the salience of the thumb among five digits, perhaps even kinesthetic ease of extension; directing it upward echoes a GOOD IS UP diagrammatic schema. Through generative sedimentation the configuration became an emblem whose current users rarely (if ever) relive these bodily motivations; they perform it because "that's what one does"—a paradigmatic case of symbolicity.

However, the process of sedimentation in other gestures contrasts with emblems like the one above, as well as with other socially shared semiotic systems such as language. Whereas language and certain social traditions predominantly involve generative sedimentation due to their rigid normativity (e.g., grammatical correctness and contextual appropriateness; Zlatev & Blomberg, 2019), gestures (again, with the exception of emblems) remain comparatively free from such strict normative constraints. Consequently, gestures significantly reflect genetic sedimentation through habitualized personal bodily experiences, allowing for greater variability and personal adaptation—even as they simultaneously engage in generative sedimentation through shared conventions.

This interplay between genetic and generative sedimentation is especially evident in recurrent gestures, which are characterized by their intermediary position between spontaneous gesticulations and emblems—an aspect Müller (2017) describes as "in-betweenness." Their forms recur across speakers and contexts, yet they also display both individual and cultural diversity (Harrison & Ladewig, 2021). Ladewig (2024) describes recurrent gestures as "sedimented individual and social practices," arguing that they "form an individual's repertoire of communicative practices," while also being "understood as a form of 'cultural action' embedded in a circle of individuation and of sustaining society" (pp. 36–37). In other words, recurrent gestures occupy a nexus where genetic and generative sedimentation intersect. Section 2.2.3 will elaborate on this point—and on recurrent gestures more generally.

In summary, sedimentation refers to the layering process through which individual (bodily) experiences solidify into durable habits (genetic sedimentation), and habitual practices accumulate through generations into socially shared structures (generative sedimentation). In so doing, sedimentation both stabilizes meaning and veils its bodily origins—a dialectical

tension that is particularly evident in gestures like the THUMBS UP but that underlies the signs of every semiotic system. Understanding this tension prepares the ground for the next section, which introduces the Motivation & Sedimentation Model that integrates embodiment, sedimented structures and situated interaction into a single explanatory framework.

2.1.4. The Motivation & Sedimentation Model (MSM)

The preceding sections showed that pan-human embodiment furnishes the primordial *motivation* for meaning, while *sedimentation* explains how such living meanings may stabilize, be forgotten, and yet continue to guide human action across historical time. What remains to be clarified is how these two forces are coordinated in actual meaning-making. Early attempts to answer this question are captured in what I will call the "simple version" of the Motivation & Sedimentation Model (MSM) (e.g., Devylder & Zlatev, 2020; Stampoulidis et al., 2019; Zlatev & Möttönen, 2023), as shown in Figure 3. Drawn on a single vertical axis, it distinguishes the Embodied level of pan-human, pre-signitive capacities and processes, the Sedimented level of conventional, community-specific norms, and the Situated level where dynamic, creative semiosis takes place. Upward solid arrows mark the motivating force of the lower levels; downward dashed arrows mark the gradual sedimentation of situated acts into more stable structures. While this simple version of the model already helped to relate the discussion of embodiment (Section 2.1.2) and sedimentation (Section 2.1.3), it leaves several issues unresolved, most notably the role of short-term, genetic (individual) sedimentation within each level and the mutual conditioning of process and structure.

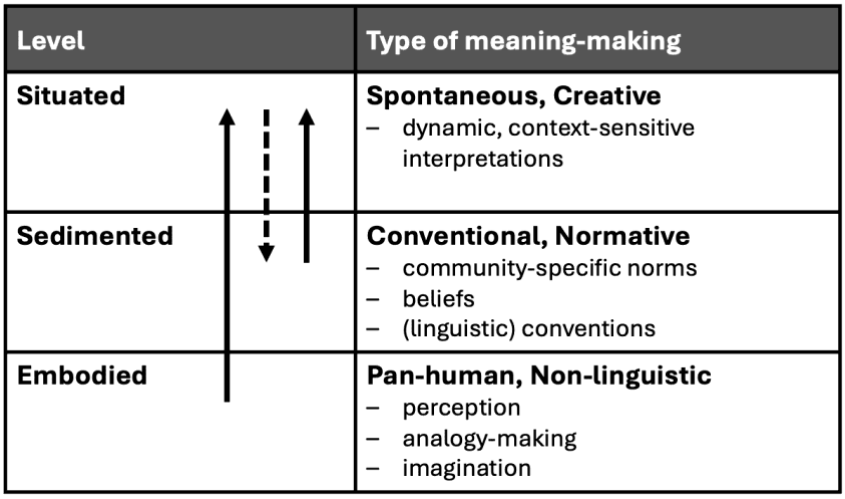


Figure 3. The "simple version" of the Motivation & Sedimentation Model (MSM) (based on Devylder & Zlatev, 2020; Stampoulidis et al., 2019; Zlatev & Möttönen, 2023)

A more elaborate answer was offered in the "complex version" of MSM introduced by Zlatev (2023), as shown in Figure 4. Here a horizontal axis was added to the original vertical one so that every level now comprises an interaction between *processes* (left column) and *structures* (right column). This horizontal axis illustrates the genetic dimension, encompassing individual-level sedimentation within shorter time scales, as outlined in Section 2.1.3. In contrast, The vertical axis in this complex model represents generative sedimentation—the historically accumulated, socially shared meanings and convention. Crucially, the interplay between motivation (depicted by solid lines) and sedimentation (represented by dotted lines) occurs along both axes, suggesting analogous dialectics between process and structure on individual and social levels (and also analogous to the two columns in the Semiotic Hierarchy; see Table 2)

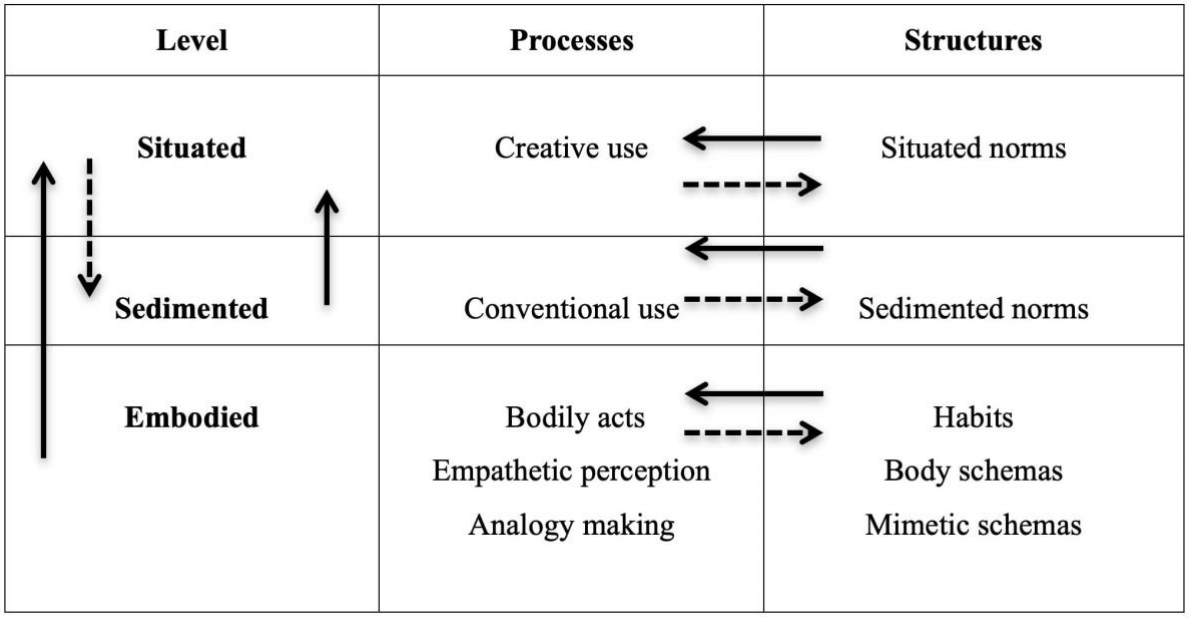


Figure 4. The "complex version" of the MSM (reprinted from Zlatev, 2023, p. 13)

However, even the complex MSM presents several theoretical inconsistencies. Firstly, distinguishing process from structure within each level does not align well with the intrinsic characteristics of each level. For instance, at the Sedimented level, processes are theoretically problematic since this level primarily encompasses stable, sedimented structures where active processes are minimally relevant. Conversely, structures at the Situated level often extend partially into the Sedimented level, undermining the neat categorization implied by the model. Another conceptual shortcoming is the orthogonal representation of genetic and generative

sedimentation. Given that these dimensions differ only in temporal *scale* and not in essence—both fundamentally concern temporality—portraying them as orthogonal may be misleading. Furthermore, the horizontal process-structure dialectic appears confined within each level, yet for example habit-formation in an individual body always already takes place in a community saturated with Sedimented norms, just as those norms themselves are continually (re)shaped by the habits of Embodied subjects. Representing (and even theorizing) genetic and generative sedimentation in terms of orthogonal relations therefore introduces the risk of obscuring precisely the kind of cross-level dynamics that the MSM set out to illuminate.

Addressing these concerns, this thesis proposes a new version of MSM as shown in Figure 5, emphasizing clear distinctions based on the characteristic features of each level. Specifically, the Embodied level strictly encompasses pan-human capacities and processes, whereas the Sedimented level exclusively concerns generatively (i.e., socially) sedimented structures whose original bodily motivations have been obscured through time. A new intermediary level, Embodied/Sedimented, captures structures arising primarily from genetic sedimentation processes—such as embodied schemas (mimetic and image schemas described in Section 2.1.2)—which inherently exhibit characteristics of embodied intersubjectivity and are thus socially shareable. In response to the reconfiguration of levels, the definition of the Situated level is broadened. Whereas previous accounts defined the Situated level narrowly as the locus of "actual social interactions" (e.g., Stampoulidis et al., 2019; Zlatev & Möttönen, 2023), the new model encompasses all forms of interaction with others and also with *the world*, including pre-social, experiential engagements. This revision addresses the conceptual fusion of genetic and generative sedimentation by situating them along a unified temporal dimension, depicted clearly by the model, which does not include the horizontal axis. Although this new version of the model introduces additional arrows to account for these revised interactions, several key points from the original MSM persist: (i) purely pan-human Embodied capacities remain intact; (ii) sedimentation from Embodied process to higher-level structures continues to be mediated through Situated interactional activities; and (iii) meaning-making at the Situated level retains its multiplicity of potential motivational sources from both (Embodied/)Sedimented structures and pre-signitive Embodied processes.

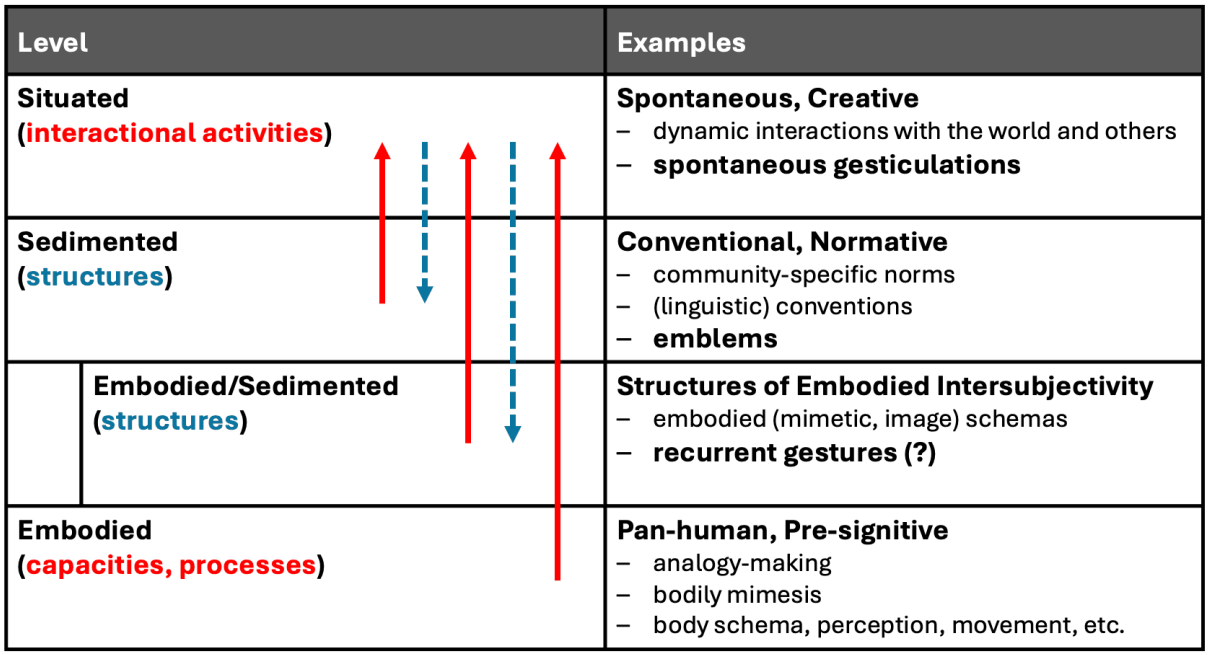


Figure 5. The revised version of the MSM, featuring the hybrid Embodied/Sedimented level between the pure Embodied and Sedimented levels: The motivation relation is represented by solid red lines and the sedimentation relation by dotted blue lines

Applying this revised MSM to the semiotic system targeted in this thesis—gesture—allows us to position distinct gestural categories clearly within the model. Spontaneous gesticulation aligns naturally with the Situated level due to its spontaneous and context-sensitive nature. Conversely, emblematic gestures, defined by social convention with faded original motivations, belong to the Sedimented level. Regarding recurrent gestures, Harrison and Ladewig (2021, p. 158, my emphasis) notably remark, "[a]s the term suggests, recurrent gestures are born through repetition and have become *sedimented forms of embodied meaning*," explicitly positioning them at the intersection of embodiment and sedimentation. Thus, recurrent gestures may be tentatively placed within the Embodied/Sedimented level, reflecting their hybrid nature. Nevertheless, the precise placement of recurrent gestures within this model and their intricate relationship between embodiment and sedimentation remains open to empirical verification through the conceptual-empirical loop central to this cognitive-semiotic thesis. More comprehensive discussions of recurrent gestures, along with spontaneous gestures and emblems, will be presented in Section 2.2.3.

In conclusion, this section has introduced, critiqued, and subsequently refined the Motivation & Sedimentation Model (MSM). This refined model not only clarifies the interrelations among embodiment, sedimentation, and situated activities but also establishes a

rigorous foundation for an empirical analysis of recurrent gestures.

2.2. Gesture

In this section, I review previous studies on gestures by organizing them from three perspectives that are particularly relevant to the current thesis: (i) the predominant semiotic ground (Section 2.2.1), (ii) the functional contributions in an utterance (Section 2.2.2), and (iii) the degree of stabilization (Section 2.2.3). It is important to note that these dimensions are not mutually exclusive. Rather, they provide different kinds of classificatory frameworks. That is, a classification in one dimension does not preclude a classification in another.

2.2.1. Gestures in cognitive semiotics

As introduced in Chapter 1, the current thesis adopts Kendon's (2004, pp. 13–14) definition of gesture based on "features of manifest deliberate expressiveness." This concept highlights the intentional and expressive nature of bodily actions that qualify as gestures. However, as Andrén (2010, 2014) points out, this definition pertains to only one aspect of gesture—namely, communicative intent (CI). In his discussion of the lower limit of gesture (i.e., distinguishing gesture from non-gesture such as instrumental action), Andrén proposes that gestures should be defined not only by their communicative intent, but also by another semiotic dimension: representational complexity (RC).

According to this analysis, CI and RC can vary independently. A bodily movement might be performed with manifest communicative intent but lack sufficient representational structure, or vice versa. What characterizes a movement as a gesture, then, is that it reaches level 3 in at least one of these two aspects—often both. CI at level 3 corresponds to Kendon's "features of manifest deliberate expressiveness," whereas RC at level 3 corresponds to the sign (function) as defined in Section 2.1.1. In other words, Level 3 in RC transforms real action into "as-if" action, where the body no longer performs an instrumental action but instead represents the action as an intentional object. Table 3 summarizes the three different levels in each of these two dimensions.

Table 3. Two separate dimensions of gestural meaning, each with three levels (reprinted from Zlatev, 2014, p. 7)

Level	Communicative intent (CI)	Level	Representational complexity (RC)
CI-3	Explicitly other-oriented action: Clear communicative intentionality	RC-3	Explicit signs: Expression X stands for meaning Y
CI-2	Action framed by mutual attunement: Unclear communicative intentionality	RC-2	Typified acts: Performance X counts as doing action-type Y
CI-1	Side effect of co-presence: No visible communicative intentionality	RC-1	Situation-specific acts: Performance X contextually suggests Y

This semiotic approach, primarily developed within the field of cognitive semiotics, further classifies gestures based on their predominant semiotic ground (Andrén, 2010; Zlatev & Andrén, 2009; Zlatev, 2014). Building on the three semiotic grounds introduced in Section 2.1.1, a gesture can be classified as:

(i) *Iconic* when the expression resembles aspects of the intentional object. Iconic gestures may depict concrete actions such as sawing (McNeill's *iconics*) or more abstract notions, for instance tracing a spiral to signify repetition (McNeill's *metaphorics*);
(ii) *Deictic* when its meaning rests on spatial contiguity and joint attention, typically realized as pointing;
(iii) *Emblematic* when the expression-intentional object (E-O) link is governed predominantly by social convention, as in the case of the THUMBS UP gesture.

As mentioned above, since semiotic grounds are not mutually exclusive, it is possible for a gesture to combine features of all three above, but one can most often be identified as the one that dominates interpretation (Zlatev, 2015b).

2.2.2. Kendon's functional classification of gestures

Kendon (2004) analyses gesture within the broader unit of *utterance*, defined as "any ensemble of action that counts for others as an attempt to 'give' information of some sort" (p. 7). As he continues, "such unit of activity may be constructed from speech or from visible bodily action or from combination of these two modalities [...]. 'Gesture' is the visible bodily action that has

a role in such units of action" (Kendon, 2004. P. 7). Accordingly, Kendon classifies gestures by *how* they contribute to the utterance's meaning, distinguishing two superordinate functions: referential and pragmatic.

On the one hand, gestures can serve *referential* functions when they "provide a *representation* of an aspect of the content of an utterance" or "by *pointing* to the object of reference in the discourse" (Kendon, 2004, p. 160, emphasis original). Such gestures are closely tied to propositional content and often co-occur with specific lexical items or phrases. Iconic and deictic gestures, as defined in the previous subsection, typically fulfill this referential function.

On the other hand, *pragmatic* functions are negatively defined, as they "relate to features of an utterance's meaning that are *not* part of its referential meaning or propositional content" (Kendon, 2004, p. 158, my emphasis). However, Kendon also provides positive definitions by explicating the three main subcategories of pragmatic functions:[4]

(i) Modal, in which gestures "indicate something about the speaker's attitude to the referential meaning or [...] contribute to the interpretative framework in terms of which this meaning should be treated" (Kendon, 2004, p. 158);
(ii) Performative, in which gestures "indicate the kind of speech act or interactional move a person is engaging in" (Kendon, 2004, p. 159);
(iii) Parsing, (or "discourse structure marker", cf. Kendon, 1995), in which gestures "make distinct different segments or components of the discourse, providing emphasis, contrast, parenthesis, and the like, or [...] mark up the discourse in relation to aspects of its structure such as theme-rheme or topical focus" (Kendon, 2017, p. 168).

Later authors retain the broad referential/pragmatic distinction while proposing alternative taxonomies of pragmatic subfunctions (e.g., López-Ozieblo, 2020); the exact subdivision is less critical for the present thesis than the superordinate referential/pragmatic distinction, which proves vital for analyzing recurrent gestures in Section 2.2.3. Recurrent gestures perform both referential and pragmatic functions (Ladewig, 2014), yet existing research has largely prioritized the pragmatic aspects (Bressem & Müller, 2014b; Payrató & Teßendorf, 2014). In

[4] Notably, Kendon (2017) later added a fourth subcategory *operational*, which is not further discussed in the present thesis.

practice, distinguishing between referential and pragmatic functions requires attention to the gesture's coordination with other components of the utterance, particularly *speech*. As Ladewig (2014) notes, "gestures with referential function are closely related to the proposition of the spoken utterance and interact with a verbal lexical unit," whereas those with pragmatic function "work on a meta-communicative level of speech" (p. 1563). Feyereisen et al. (1988) likewise emphasize the gesture-speech relation as a diagnostic criterion. One way of stating this is that referential gestures function as *elements* that constitute an utterance, operating at the word or phrase level in speech, whereas pragmatic gestures are "deployed meta-communicatively and *operate upon a speaker's utterance*" (Ladewig, 2014, p. 1563, my emphasis).

Then, how does this referential/pragmatic distinction, along with the criteria for distinguishing the two, relate to the concept of the sign in cognitive semiotics, defined by the signification relationship between expression (E) and intentional object (O) introduced in Section 2.1.1? Mouratidou et al. (2024) argue that bodily expressions such as gestures, when they qualify as signs, possess "denotational" meaning that is categorized based on the predominant semiotic ground, as described in Section 2.2.1. In addition to this, bodily expressions may also carry "non-denotational" meaning—"expressing emphasis, modality (uncertainty, rejection, etc.), and affect (surprise, repulsion, etc.)" (Mouratidou et al., 2024, p. 238). This distinction to some degree corresponds to the referential/pragmatic distinction used by Kendon and his followers. However, the term "denotation" is too narrow for the relation between E and O, and hence in this thesis I will use the more general notion of *signification*. In short, in sign use, E signifies O.

Furthermore, Ladewig's (2014) above criteria can be reinterpreted in cognitive semiotic terms as follows: signs can have intentional objects that are discrete and specific, in the sense that they function as *parts* within a larger whole. Such wholes may include an utterance (in Kendon's sense), a proposition, or a situation. I will refer to these as *delineated* intentional objects (meanings). In contrast there are other kinds of intentional objects that rather serve as *qualifications* of a whole, attributing certain qualities that shape how the whole is experienced or interpreted. These will be referred to as *holistic* intentional objects (meanings).

Because this thesis analyzes gestures from the cognitive-semiotic perspective, this phenomenologically defined concept of the sign will be adopted, with two kinds of intentional objects and their corresponding meanings: *delineated* and *holistic.* For example, in response to the question *Which sports does he like?*, if one answers *I think he likes baseball* while enacting a throwing gesture, the gesture signifies a delineated intentional object—THROWING A BALL—

which constitutes *part* of the whole utterance. With such delineated meanings, gestures often provide redundant or complementary information to the spoken component (Ladewig, 2014)—in this case, contributing to the reference to baseball. In contrast, if one performs a shrugging movement alongside the same speech, perhaps accompanied by additional cues like rising intonation or lexical hedges such as *or something*, the shrugging gesture signifies a holistic intentional object of UNCERTAINTY. In this case, the gesture *qualifies* the whole utterance as an expression of the speaker's epistemic stance.

2.2.3. Recurrent gestures

As recurrent gestures have already been discussed several times in the preceding sections, it can be assumed that readers have acquired some initial understanding of what they are. This section aims to reinforce that understanding by offering a more detailed account of the defining characteristics of recurrent gestures and situating them within a continuum from iconicity (and idiosyncrasy) to symbolicity (and conventionality).

As noted in Section 2.1.3, recurrent gestures may be seen as occupying the nexus between genetic (individual) sedimentation and generative (social) sedimentation. In line with this, Müller (2017, p. 280, my emphasis) defines recurrent gestures as "conventionalized gestural expressions that have a basic form and a prototypical meaning *within a given speech-gesture community and also within an individual speaker.*" Because Cooperrider (2019) places recurrent gestures and emblems in the broader category of what he calls "gestural conventions," they resemble emblems in exhibiting "relatively constrained mappings between specific meanings and specific forms" (p. 215). They differ, however, in that recurrent gestures comprise "a broader range of interrelated forms and meanings" (Cooperrider, 2019, p. 216) than emblems. According to Müller and her colleagues' form-based approach to (recurrent) gestures (e.g., Müller et al., 2013), which begins with "a detailed analysis of gesture form—both regarding their articulatory (etic) and their meaningful (emic) features or clusters of features—as a point of departure to reconstruct meaning" (Müller, 2014, pp. 138–139), recurrent gesture research centers on what she terms the "kinesic core." The kinesic core of a given recurrent gesture is defined as "a shared Gestalt of selected features or parameters (for instance, hand-shape and orientation) that does not vary across contexts and that comes with a more or less conventionalized basic prototypical meaning" (Müller, 2017, p. 280). Despite the conventionalized nature of the kinesic core, the remaining formal features "may be used to specify and alter the meaning of the kinesic core spontaneously and according to local

affordances of the communicative situation" (Müller, 2017, p. 280), making recurrent gestures also similar to spontaneous, idiosyncratic gesticulations. Thus, recurrent gestures are "hybrids of idiosyncratic and conventional elements," occupying an intermediate position between spontaneous gestures and emblems in the continuum of increasing conventionalization (see Figure 6).

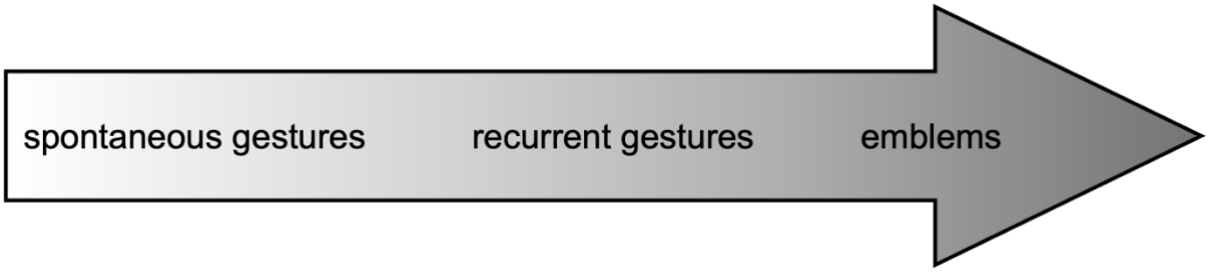


Figure 6. A continuum of gestures from iconicity (and idiosyncrasy) to symbolicity (and conventionality). (adapted from Ladewig, 2024, p. 42)

Shifting focus from the sedimented nature of recurrent gestures (see also Section 2.1.3) to their embodied basis, what do Harrison and Ladewig (2021) mean when they describe recurrent gestures as "sedimented forms of *embodied* meaning" (p. 158, my emphasis)? Put simply, as numerous studies on (recurrent) gestures have argued (e.g., Bressem & Müller, 2017; Mittelberg, 2017; Mittelberg & Joue, 2017), recurrent gestures are grounded in instrumental action. Harrison and Ladewig (2021, p. 156) note that recurrent gestural forms and their corresponding prototypical meanings are "derived from (and motivated by) manual actions involved in the everyday manipulation of objects, such as gripping, offering, and sweeping away." This is especially evident in the case of *Away gesture families* (Bressem & Müller, 2014a), which are associated with negation and refusal (Harrison, 2014). Building on Fillmore's (1982) notion of semantic frames, Bressem and Müller (2017) argue that both the delineated and holistic intentional objects or meanings, in the terms of the present thesis, of Away gestures are grounded in what they term the "away action scheme," an experiential frame consisting of indexically connected elements:

> The elements included in this action scheme are an unpleasant situation (starting point) in which annoying objects are in the immediate surrounding (cause). These are removed through an action of the hand (action), which then leads to removal of the objects and a neutral situation (endpoint). (Bressem & Müller, 2017, p. 6)

Notably, this rich experiential structure would correspond to the notion of *mimetic schema* introduced in Section 2.1.2. When such an action scheme or mimetic schema is re-enacted, the resulting Away recurrent gesture can signify a delineated meaning (profiling the action itself, e.g., brushing dust off a shelf) or a holistic meaning (profiling the action's goal, e.g., dismissing an unwelcome topic).

However, the embodied grounding in manual action is only one of the two primary motivations for the recurrent kinesic core described by Müller (2017). As she observes, "there appear to be (at least) two different ways in which a recurrent kinesic core may be motivated through embodied experiences: schematic enactment of an instrumental action or of *a movement pattern*" (p. 294, my emphasis). The latter motivation is illustrated by cyclic gestures (Ladewig, 2011; Ruth-Hirrel, 2018), which are grounded in "sensory-motor experiences of cyclic movements of the hands and arms" and are re-enacted to signify "repetition, cycle, duration, or continuation" (Müller, 2017, p. 293). Unlike mimetic schemas, this second type of motivation is better understood through the lens of image schemas, also introduced in Section 2.1.2. In her cognitive-linguistic analysis of cyclic gestures, Ladewig (2011) draws on the CYCLE image schema and its metaphorical projections to abstract domains through the conceptual metaphors TIME IS MOTION THROUGH SPACE and MIND IS A MACHINE (Lakoff & Johnson, 1980). In her spoken corpus of 56 instances of cyclic gestures in German, 37 tokens were used to express holistic meanings related to word or concept search. Analogous to a computer's loading spinner, Ladewig (2011, p. 8) argues, "the mental activity of searching, which can be considered as an outcome of the working mind, is conceptualized as a circuit." In her account, this conceptualization involves the projection of the CYCLE image schema onto the abstract domain of MIND via the MIND IS A MACHINE conceptual metaphor (Ladewig, 2011). Taken together, these possible embodied motivations for recurrent gestures are in line with the broader argument that both image and mimetic schemas play an important role in gestural meaning-making (Cienki, 2013).

Given their dual grounding in embodied experiences and (genetic and generative) sedimentation, recurrent gestures have been described by Cooperrider (2019) as "natural conventions," meaning that "they are culturally selected (i.e., conventionalized) from a menu of motivated (i.e., natural) options" (pp. 228–229). As a consequence of this proposal, he points out two important implications: "First, there will be very few absolute universals—specific recurrent gestures or gestural practices that are found the world over. But, second, there will be

very few one-off cases that are found in one place and only one place" (Cooperrider, 2019, p. 229). To empirically test this claim, Ladewig (2024) emphasizes the importance of conducting cross-cultural studies of recurrent gestures, explicitly stating that "cross-cultural comparisons can spell out the continuum from 'one-off cases' to absolute gestural universal" (p. 48). Yet, there have been few such comparative studies. One is Bressem and Wegener's (2021) cross-linguistic analysis of the discourse-structuring function of *Holding Away gesture* in German and Savosavo. Their study revealed that while both groups employed the gesture to mark contrast and conclusion, Savosavo speakers uniquely used it for elaboration, whereas only German speakers employed it to indicate inference. Crucially, the authors attribute these differences not primarily to linguistic or cultural differences, but to methodological variables such as recording purposes (e.g., for research vs. entertainment), interactional formats (monologic vs. dialogic), and genre-based discourse expectations (Harrison et al, 2021; Ladewig, 2024). Their conclusion underscores the importance of using comparable data collection methods when conducting cross-cultural gesture research, a point this thesis fully acknowledges and addresses. In this light, the current thesis examines how the Two-Handed-Alternation-Sagittal (2HAS) gesture (see the next section for more details) is used in Japanese and Swedish—two linguistically and culturally distinct communities—under comparable methodological conditions. By doing so, it aims to empirically test Cooperrider's (2019) claim that neither absolute universals nor one-off cases dominate, and to shed light on how recurrent gestures reflect both pan-human embodiment and genetic/generative sedimentation, as evidenced by their universality and individual/cultural diversity.

2.3. Target gesture: Two-Handed-Alternation-Sagittal (2HAS) gesture

The Two-Handed-Alternation gesture on the Sagittal axis (2HAS)—the target recurrent gesture in this thesis—emerged in my previous studies as a gesture that plays a significant role in forming what Construction Grammarians (e.g., Goldberg, 1995; Hilpert, 2019) refer to as a "multimodal construction" (cf. Zima & Bergs, 2017): a conventionalized polysemiotic sign complex consisting of two or more expressions that signify the same or related intentional objects, in cognitive-semiotic terms.[5] By reviewing my previous studies on multimodal

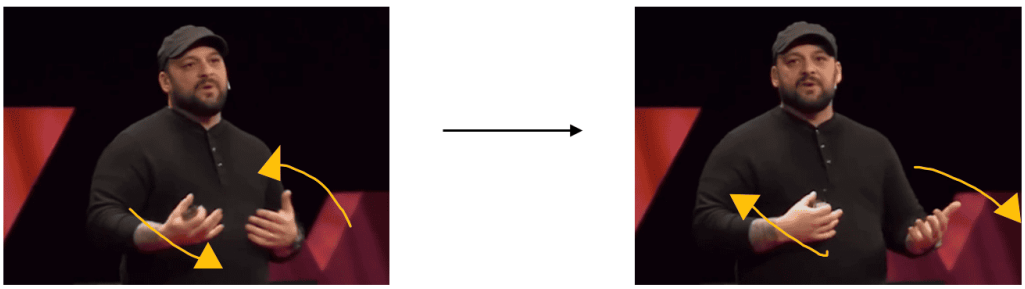
[5] Researchers in cognitive semiotics (Zlatev & Möttönen, 2023) distinguish between *multimodality*, defined as the combination of multiple sensory-motor modalities, and *polysemiosis*, defined as the combination of multiple semiotic systems (e.g., language, gesture, depiction). Therefore, although the

constructions, this section provides evidence that the 2HAS gesture qualifies as a recurrent gesture (Section 2.3.1) and that it is grounded in pan-human embodied experiences of manual actions and movement patterns (Section 2.3.2).

2.3.1. 2HAS as a recurrent gesture

Kuryu (2025) used the TED Corpus Search Engine (TCSE; https://yohasebe.com/tcse/) to investigate whether a set of formally and semantically similar constructions were used significantly frequently with their gestural correlates to form multimodal constructions. The target constructions included *back and forth*, *up and down* and *in and out*, all of which were frequently accompanied by gestures with bidirectional movement. To test whether the combinations of the target constructions and their frequent counterparts in these bidirectional gestures could qualify as multimodal constructions, I employed a statistical method known as "crossmodal collostructional analysis" (Uhrig, 2021). This method goes beyond mere co-occurrence frequency by measuring the *salience* of a gesture for a given construction—specifically, by comparing how frequently the gesture appears with the construction to how often it occurs in a more neutral reference dataset. The premise is that the more frequently a gesture co-occurs with a specific construction and the less frequently it appears in the reference dataset, the more salient it is for that construction. According to the criteria of (i) mere frequency and (ii) salience (as measured by crossmodal collostructional analysis), the 2HAS gesture emerged as a likely gestural component in multimodal constructions involving *back and forth* (N = 106) and *in and out* (N = 49). The 2HAS gesture accounted for 24% (N = 26) of gestures co-occurring with *back and forth*, and 20% (N = 10) with *in and out*. In addition to these findings—which suggest that 2HAS is at least locally conventionalized with these particular constructions—it is also worth noting that the gesture appeared with substantial frequency (N = 11) in the reference dataset of 560 gestures observed in more neutral contexts. Figure 7 presents the description of the 2HAS gesture as found in Kuryu (2025).

term "multimodal construction" as used in cognitive linguistics actually refers to the combination of multiple semiotic systems—and would thus be more appropriately termed "polysemiotic constructions"—this thesis will retain the term "multimodal construction" due to its greater conventionality.



Description	Both hands alternate quickly in a back and forth movement on the sagittal axis.
Features	2H, HS: open/forefinger, Loc: in front of torso Move: typically movement from the wrist
Frequency	OVER: 5 (3.3%), AGAIN: 1 (0.9%), ON: 1 (2.5%) BF: 26 (24.5%), UD: 2 (2.4%), IO: 10 (20.8%) Reference data: 11 (2.0%)

Figure 7. Typical formal features of the 2HAS gesture found in Kuryu (2025)

Based on the finding in Kuryu (2025) that bidirectional gestures across various axes were also frequently observed with *back and forth*, a follow-up study (Kuryu, 2024a) further analyzed the formal features of bidirectional gestures in general. Using the NewsScape 2016 gesture-annotated corpus of Red Hen Lab (https://www.redhenlab.org), 215 bidirectional gestures accompanying *back and forth* were annotated in terms of seven formal features, of which articulator (1-hand, 2-hand, head, body), movement type (applicable only to 2-hand gestures: parallel [QR1], alternation [QR2], other) and movement axis (vertical, horizontal, sagittal) were used to categorize movement patterns. Table 4 shows the distribution of dominant movement types (N = 186) by the movement axis, which was significant ($X^2 = 98.277$, df = 4, $p < 2.2e\text{-}16$). A Pearson residual post-hoc test confirmed that cells highlighted in red occurred significantly more frequently than expected, while those in blue were significantly less frequent.

QR1[6] QR2[7]

[6] http://go.redhenlab.org/dak/0003
[7] http://go.redhenlab.org/dak/0008

Table 4. Distribution of movement patterns by movement axis (kin*etic* grid)

	1H	2H-parallel	2H-alternation	TOTAL
Vertical	3	4	**13**	20 (11%)
Horizontal	**66**	**37**	0	103 (55%)
Sagittal	**24**	0	**39**	63 (34%)
TOTAL	93	41	52	186

Among the possible realizations of bidirectional gestures on each axis, the 2HAS gesture in particular qualifies as a recurrent gesture, as Table 4 supports its conventionalized nature in two respects: (i) on the sagittal axis, 2HAS occurred more frequently than other anatomically possible variants (notably, compare 2HAS with 2H-Parallel-Sagittal); and (ii) despite its greater articulatory effort, 2HAS was more frequent than 1H-Sagittal. This contrasts with bidirectional gestures on the horizontal axis, which display a more anatomically driven tendency: (i) 2H-Alternation-Horizontal was not used at all, likely due to its awkwardness; and (ii) 1H-Horizontal was more frequent than the 2H-Parallel-Horizontal, likely because the latter requires greater articulatory effort.[8] Notably, the 2HAS gesture is also documented in Bressem & Müller's (2014b, p. 1580) list of German recurrent gestures—coincidently under the name of "back and forth"—and is claimed to have the semantic core of CHANGE, UNCERTAINTY and AMBIVALENCE. Although Bressem and Müller (2014b) do not report the frequency of the 2HAS gesture in their corpus, Ladewig (2025) found 27 instances of 2HAS in her five-hour corpus of German political talk shows.

2.3.2. Different embodied motivations for the single kinetic pattern of 2HAS

From a more qualitative perspective, Kuryu (2024b) demonstrates that there are two different embodied motivations for the multimodal construction [*back and forth* + 2HAS], and more generally, for the recurrent gesture 2HAS itself. This difference in motivation becomes particularly evident when focusing on the type of verbs that precede *back and forth* (i.e., intransitive vs. transitive). As discussed in Section 2.2.3, Müller (2017, p. 294) argues that there

[8] Still, it is possible that the 2HAS is only locally conventionalized as a component of the [*back and forth* + 2HAS] multimodal construction. This may be due to factors such as the incompatibility of other bidirectional gestures on the sagittal axis—which typically begin with a *forward* movement—with the articulation of '*back* and forth'. In contrast, the 2HAS inherently involves both forward and *back*ward movement from the outset.

are two different embodied motivations for a recurrent kinesic pattern: "schematic enactment of an instrumental action or of a movement pattern." When combined with intransitive verbs (e.g. *coming* [QR3], *going* [QR4], *flying* [QR5]), the multimodal construction is likely motivated by a movement pattern of BIDIRECTIONALITY. However, there are cases where the 2HAS appears to be grounded in instrumental actions, such as when it accompanies the verb *exchange* (QR6 and QR7), with the speaker schematically re-enacting the instrumental actions GIVE and RECEIVE. This is also the case when the multimodal construction is used with transitive verbs (e.g. *sending NP* [QR8], *throwing NP* [QR9]). Notably, these two different embodied motivations for the multimodal construction (and thus 2HAS)—movement patterns (BIDIRECTIONALITY) and instrumental actions (GIVE & RECEIVE)—seem to correspond to the concepts of image schemas and mimetic schemas, respectively (see Section 2.1.2).

QR3[9] QR4[10] QR5[11] QR6[12] QR7[13] QR8[14] QR9[15] QR10[16]

Regarding the role of these embodied schemas in the construal of an intentional object shaped by a gestural expression, Cienki (2015, p. 509, my emphasis) notes:

> Whereas the level of specificity relates most clearly to lexical choice (e.g., finger vs. hand vs. forearm), gesture can provide additional (visual) information in this regard about how referents are to be conceptualized, for example *by virtue of the degree of schematicity in the form of the gestures used*, be it on the more simple schematic level (such as that of image schemas à la Johnson, 1987, like PATH or CYCLE) or closer to basic-level human actions (like those discussed in Zlatev, 2005, as mimetic schemas, like PUT IN or RUN).

For instance, consider examples QR8 and QR9 where transitive verbs precede *back and forth*,

[9] http://go.redhenlab.org/dak/0034
[10] http://go.redhenlab.org/dak/0009
[11] https://www.youtube.com/watch?v=aiIxdFBA0Sw&t=323s
[12] https://www.youtube.com/watch?v=BAswj8evFZk&t=74s
[13] https://www.youtube.com/watch?v=HuWMO3M1HVQ&t=88s
[14] https://www.youtube.com/watch?v=SU8JYKGekXo&t=465s
[15] http://go.redhenlab.org/dak/0026
[16] http://go.redhenlab.org/dak/0020

in comparison to QR10, where the speaker discusses the act of *sharing* something. Lexically, all three examples are similar in terms of the delineated intentional objects signified by 2HAS, as they all concern EXCHANGE of (virtual) objects schematically (i.e., the enactment of the GIVE and RECEIVE mimetic schemas). However, in QR10, the speaker extends his index fingers while performing 2HAS—a variation that likely evokes the image-schematic structure of BIDIRECTIONALITY in the perceiver's mind, rather than the mimetic schemas. This subtle variation in the gestural expression results in a different construal: instead of profiling the action of EXCHANGE itself, the gesture highlights the BIDIRECTIONAL aspect of that action. This distinction will be particularly relevant in the upcoming analysis of the delineated meanings of 2HAS as performed by Japanese and Swedish speakers.

2.4. Research aims and theoretical expectations

Building on the theoretical scaffolding established in Sections 2.1–2.3, this section clarifies the concrete aims that lead to the forthcoming empirical investigation. In particular, it specifies how the two research questions introduced at the end of Chapter 1 are expected to play out once the Two-Handed-Alternation-Sagittal gesture (2HAS) is examined across Japanese and Swedish speakers.

Regarding RQ1, as stated in Section 1, and elaborated in Section 2.3.2, considering that the embodied motivations for 2HAS lie in *pan-human* bodily experiences of manual actions (GIVE & RECEIVE) and movement patterns (BIDIRECTIONALITY), it can be expected that, when these embodied schemas are enacted, the delineated meaning of 2HAS would exhibit commonality between Japanese and Swedish speakers. However, with regard to the inclusion of UNCERTAINTY in the semantic core of 2HAS in Bressem and Müller's (2014b, p. 1580) formulation, this intentional object of UNCERTAINTY, as illustrated by the shrugging example above, is likely to manifest as a holistic meaning—specifically, as a speaker's epistemic stance toward the utterance as a whole; that is, as "a *qualification* of a whole." In such cases, it becomes difficult to identify any direct connection between this holistic meaning and embodied experiences such as manual actions like GIVE & RECEIVE or movement patterns like BIDIRECTIONALITY. In other words, this potential holistic meaning derived from the German corpus used by Müller and Bressem (2014b) is likely not based on embodied experience but rather a result of generative (social) sedimentation, in which the original motivation has been forgotten. Therefore, in general, compared to delineated meanings, holistic meanings—where embodied motivations are less transparent—can be expected to show greater variation between

Swedish (a language closely related to German) and Japanese, which differs significantly in both cultural and linguistic background.

This leads naturally to RQ2. In Section 2.1.4, recurrent gestures were tentatively situated at the hybrid Embodied/Sedimented level of the MSM, based on prior literature on recurrent gestures. Given that the delineated meaning of 2HAS is grounded in embodied schemas as structures of embodied intersubjectivity, this positioning appears to be appropriate. Here, it is also important to emphasize the distinction between genetic and generative sedimentation. As noted in Section 2.1.4, structures of embodied intersubjectivity at the Embodied/Sedimented level are primarily formed through genetic (individual) sedimentation. These structures are not entirely detached from their embodied motivations and thus such motivations may be relatively more accessible and recoverable than those based on generative sedimentation. In contrast, the motivation for the potential holistic meaning of 2HAS, such as the UNCERTAINTY predicted at least for the Swedish group, and more broadly, structures located at the pure Sedimented level (such as linguistic expressions), are generally products of generative sedimentation, where the original embodied motivation is significantly lost. Therefore, it is quite possible that even within a single gesture like 2HAS, the source of the primary motivation may differ depending on the type of meaning—delineated or holistic—whether it stems from the Embodied/Sedimented level as a product of genetic sedimentation or from the pure Sedimented level as a result of generative sedimentation. Moreover, this potential difference in motivational origin—genetic vs. generative sedimentation—could also be inferred from other factors beyond the meaning of 2HAS itself, such as differences in the frequency of 2HAS across cultural groups or among individuals within each group. Accordingly, such factors will also be taken into consideration in the analysis.

In sum, we may expect (i) *cross-cultural convergence* in delineated meanings of 2HAS grounded in pan-human embodied schemas, but (ii) *culturally contingent divergence* in holistic meanings that have drifted into the realm of sedimented convention. By testing these expectations as described in the following chapters, I aim to provide empirical insight into the dynamic interplay between embodiment, sedimentation, and recurrency in gestural practices.

Chapter 3. Methods

This chapter explains the analytical procedure based on phenomenological triangulation, integrating first-, second-, and third-person perspectives to examine how the 2HAS gesture reflects both embodied experience and sedimented practices (Section 3.1), and outlines the empirical methodology employed to investigate how the 2HAS gesture operates as a recurrent gesture within and across two cultural contexts (Section 3.2).

3.1. Application of phenomenological triangulation

This section details how the methodological principle of phenomenological triangulation (Section 2.1) was applied to the present study. As summarized in Table 5, each perspective contributed distinct but complementary insights. The following subsections explain how each was implemented—from structuring the design of this study to analyzing the 2HAS gesture, while Section 3.2 elaborates all the necessary details.

Table 5. Phenomenological triangulation as applied in the present thesis (see Section 2.1)

Perspective	Methods	Applied in this study
First-person ("subjective")	Conceptual analysis Phenomenological methods Systematic intuitions	Analysis of 2HAS motivations Shaping RQs and hypotheses Designing methodological part
Second-person ("intersubjective")	Empathy Imaginative projection	Literature review Interpreting the meanings of 2HAS Inter-coder validation
Third-person ("objective")	Detached observation Experimentation Computational modelling	Quantification of results Beysian Poisson regression analysis

3.1.1. First-person perspective

This study is rooted in my phenomenological reflection on recurrent gestures, especially 2HAS. It was through repeated personal experience with this gesture—its use, observation, and interpretation—that the initial intuitions and hypotheses were formed. These reflections did not serve merely as inspiration but as the basis for a systematic first-person method, aligning with the core principles of cognitive semiotics as outlined in Chapter 2. The generalizations about the embodied motivations identified for 2HAS, namely the schematized instrumental actions of GIVE & RECEIVE and the visuo-spatial pattern of BIDIRECTIONALITY, emerged from these

reflections. They became essential for grounding the theoretical expectations and empirical design of this research. These embodied experiences, analyzed in Section 2.3.2, are not culturally specific but rather rooted in pan-human bodily interaction with the world. This realization prompted the central premise behind the cross-cultural design of the study: if the embodied motivations behind 2HAS are indeed universal, then the gesture should appear, with comparable form and meaning, in speech communities as culturally and linguistically distinct as Japan and Sweden. This line of reasoning led me to predict a degree of cross-cultural convergence in the delineated meanings of 2HAS, as such meanings are more transparently linked to the bodily motivations of BIDIRECTIONALITY and GIVE & RECEIVE schemas.

However, this reflection also revealed a significant layer of complexity. In their work on German recurrent gestures, Bressem and Müller (2014b) identified UNCERTAINTY as part of the semantic core of 2HAS. Drawing on my intuitions about English and Japanese, I noticed that English speakers tend to use 2HAS in a manner analogous to lexical hedges such as *kind of*—serving as an epistemic marker indicating uncertainty or softening a claim. In contrast, Japanese speakers, myself included, are more likely to use 2HAS to mark the process of word or concept search. Because these holistic meanings—epistemic marking or word/concept search—are difficult to trace back to the original embodied motivations of 2HAS, I inferred that they likely stem from social or discursive conventions, i.e., from generatively sedimented practices rather than embodied schemas. From this perspective, such meanings are less likely to exhibit cross-cultural uniformity and may diverge between groups that differ in their cultural proximity to English, such as Swedish and Japanese. Thus, I hypothesized that while delineated meanings of 2HAS would remain largely consistent across both groups, holistic meanings would vary, reflecting generative sedimentation in each culture.

Consequently, to test these expectations directly, I developed the comparative study design outlined in Section 3.2, using matched interactional settings and conversational topics. This methodological structure allows for an empirical investigation of how embodiment and sedimentation interact in shaping the recurrent use of gestures—directly addressing the core concern of RQ2.

3.1.2. Second-person perspective

Many steps in the study also required understanding another person's point of view—that is, adopting a second-person perspective. While the first-person conceptual analysis laid the foundation for the study's hypotheses, it was itself dependent on the reading and interpretation

of others' research—a fundamentally intersubjective process. The literature review presented in Chapter 2 involved a dialogical engagement with the intentions, arguments, and terminologies of other scholars, thus already implicating an intersubjective mode of understanding at the conceptual level.

A second-person perspective also played a critical role in the annotation phase that is described in Section 3.2.5. In analyzing the meanings of individual instances of the 2HAS gesture performed by participants, I relied heavily on empathy (e.g., Zahavi, 2018)—striving to understand the gestures not merely as externalized representations, but as intentional acts performed by embodied subjects. This required attempting to inhabit the perspective of each participant, discerning not only what was said, but how it was meant and experienced by the speaker. The annotation process, then, was not a mere classification of data but an interpretive act of intersubjective resonance.

Moreover, this second-person stance extended beyond the participant-researcher relationship to include collaboration between researchers. For the Swedish data, the research partner was a native speaker; for the Japanese data, I served in that role. Once each of us had completed our annotations, we revisited cases in which one of us had attributed a meaning unfamiliar to the other's linguistic intuition. This reciprocal reinterpretation required each coder to adopt, however provisionally, the communicative stance of both a foreign language user and a fellow analyst, thereby deepening the intersubjective validation of every coding decision. Finally, discussions with my supervisor on the theoretical expectations and methodological design—as well as on key concepts, including the distinction between delineated and holistic intentional objects, which is novel for cognitive semiotics—also formed a second-person aspect.

3.1.3. Third-person perspective

Finally, the third-person perspective is manifested in the more detached phases of this research, particularly in the formal annotation and statistical analysis of gesture data. While meaning annotation required intersubjective interpretation, the annotation of the formal features of 2HAS—such as hand configuration, movement axis, and gesture position—was carried out using Bressem's (2013) notation scheme. These physical parameters of the gesture were treated independently of their contextual interpretations, thereby exemplifying a third-person approach grounded in observable, measurable aspects of the gesture.

Furthermore, the empirical data presented in Chapter 4 are expressed through quantified

aggregates of these observed gesture instances. Such aggregation entails a move away from individual subjective experiences to more generalizable patterns, allowing for inferential statistical modeling. Specifically, the use of Bayesian negative-binomial regression analysis (Winter & Bürkner, 2021) constitutes a third-person method par excellence, offering a probabilistic account of gesture distributions across groups. This analytical stance reflects a form of epistemic detachment from "the things themselves" (to use Husserl's famous phrase), thus complementing the more phenomenologically engaged first- and second-person methods.

3.2 Design of the study

The study was designed to elicit naturalistic, spontaneous conversational data from native speakers of Japanese and Swedish. Participants were recorded in dyadic interactions, discussing a range of everyday topics—half of which were selected to encourage the use of 2HAS, while the other half served as controls. Care was taken to maintain consistency across the two language groups in terms of task structure, topic content, and recording conditions. This design enables a comparative analysis of gesture use that reflects both within-group and cross-cultural patterns.

3.2.1. Participants

Regarding the Japanese group, 20 native Japanese speakers (12 women, 8 men) participated in the study, ranging in age from 19 to 23 (median: 20). They were recruited via social media and personal acquaintances, with most being exchange students from Japanese universities. During data collection sessions, they were asked to work in pairs, resulting in a total of 10 dyads (2 male-female, 5 female-female, and 3 male-male pairs).

As for the Swedish group, 20 native Swedish speakers (10 women, 10 men) took part in the study, with ages ranging from 20 to 53 (median: 23). They were recruited through social media and personal connections of mine and the Swedish research partner. As with the Japanese participants, they worked in pairs during data collection, forming a total of 10 dyads (4 male-female, 3 female-female, and 3 male-male pairs). All participants were at the time living in Southern Sweden, where the study was conducted.

Participants were also asked about their proficiency in the other language (either Japanese or Swedish) and how well they knew their conversation partner. All participants reported a proficiency below the beginner level (corresponding to A1 or A2 on the CEFR scale), and each

conversational pair had known and interacted with each other prior to the data collection session. Table 6 summarizes the relevant participant information.

Table 6. Participant characteristics

	Age			Gender		Proficiency of the other language	
Group	**Range**	**Average**	**Median**	**Male**	**Female**	**Little or no knowledge**	**Beginner (A1 or A2)**
Japanese	19-23	20.6	20	8	12	14	6
Swedish	20-53	24.7	23	10	10	14	6

3.2.2. Materials

During the data collection sessions, each pair was asked to discuss questions related to six everyday topics: (1) communication, (2) travel and transportation, (3) different kinds of relationships, (4) changes, (5) habits and routines, and (6) happiness (see Appendix A). The first three topics—communication, travel and transportation, and relationships—were selected as 2HAS-prompting topics based on my reflections on when 2HAS gestures are typically used in English conversations (see Section 3.1). This intuition is supported by the attestation of 2HAS gestures accompanying speech related to these topics, as illustrated in the example videos accessible via QR codes 11 (communication), 3 and 5 (travel and transportation), and 12 (relationships). In accordance with the analysis of 2HAS presented in Section 2.3.2, the 2HAS gestures accompanying communication-related speech are regarded as grounded in basic manual actions GIVE and RECEIVE; while those accompanying speech about travel/transportation and relationships are seen as based on the more schematic structure of BIDIRECTIONALITY. The other three general topics—changes, habit and routines, and happiness—were chosen as controls, as 2HAS gestures were not expected to be the most fitting gesture to use when discussing these topics.[17] Each of the six topics includes three questions, resulting in a total of 18 questions (see Appendix A).

[17] While Bressem and Müller (2014b, p. 1580) include CHANGE as a semantic core of 2HAS, the study treat it as a non-2HAS-prompting topic, as there seems no association between CHANGE and 2HAS's motivating forces; either BIDIRECTIONALITY or GIVE and RECEIVE.

QR11[18] QR12[19]

Before data collection, the questions were translated into Japanese and Swedish as closely as possible by myself and my Swedish research partner, both native speakers of their respective languages. The translated versions were saved as separate PDF files for use during the data collection sessions (see Appendices B and C).

3.2.3. Procedure

During the data collection sessions, each pair of participants first completed a survey covering the information summarized in Table 5. They were then asked to discuss questions related to six everyday topics in their native languages, with the conversations video-recorded. My role, along with that of the Swedish research partner, was to provide the questions and the same set of example answers (see Appendix A) when participants had difficulty coming up with ideas. We also managed the time to ensure that the entire recording lasted approximately 48 minutes, with 8 minutes allocated to each topic. No specific time limit was set per question, resulting in some pairs discussing all three questions within a topic and others only one or two. This was intended to prioritize the naturalness of the conversations.

In each data collection session, the six topics were randomized so that 2HAS-prompting topics and non-2HAS-prompting topics alternated. Additionally, the order in which the topics were presented was arranged so that odd-numbered pairs started with a 2HAS-prompting topic, while even-numbered pairs started with a non-2HAS-prompting topic (see also Appendix A).

3.2.4. Data collection, data management and ethics

Data collection was conducted in a room at the Center for Languages and Literature at Lund University between December 2024 and March 2025. Before each session, participants received a consent form outlining the study's purpose and procedures. While the form noted

[18] https://www.youtube.com/watch?v=PrUA8L40Dic&t=44s
[19] https://www.youtube.com/watch?v=pLqjQ55tz-U&t=779s

that the study's focus included manual movements, it did not specifically mention the 2HAS gesture. All participants provided informed consent and were made aware of their right to withdraw at any time during or after the session (see Appendix D). Each participant was assigned an ID, and all information related to them (such as the details in Table 6) was anonymized. No sensitive information—such as ethnic origin, political views, religious beliefs, or sexual orientation—was collected.

Except for one Swedish pair, whose conversation lasted 44 minutes, all other 19 pairs spoke for approximately 48 minutes. All video recordings are securely stored on offline storage devices. After each data collection session, participants received a movie ticket.

3.2.5. Data annotation

As the total amount of data is substantially large (ca. 16 hours), only candidate 2HAS gestures were annotated using ELAN (ELAN, 2022). As discussed in Section 2.3.1, while the formal features listed in Figure 7 may be considered typical, the core defining components of 2HAS gestures—i.e., the kinesic core (see Section 2.2.3)—are the specification of handedness (2H), movement type (alternation), and movement axis (sagittal), as the name Two-Handed-Alternation-Sagittal gesture suggests. As a first step, all kinetic patterns exhibiting these essential features were coded as candidate 2HAS gestures, without considering speech content (i.e., with the audio muted). At this stage, movement phase and phrase annotation was conducted following Kita et al.'s (1998) annotation scheme. The candidate 2HAS gestures were then further annotated for additional formal features such as pivot (either wrist or elbow), gesture position (McNeill, 1992), hand shape, hand orientation (Bressem 2013), and action phase (segments within a stroke; Hinnell, 2018). Importantly, these formal features were used to help interpret the meanings of 2HAS gestures in the next step, although a comprehensive analysis of these features is beyond the scope of this thesis.

In the second step, the meanings of 2HAS gestures were analyzed in relation to the speech with which they co-occurred. Here the gestures' meanings were examined with attention to their temporal alignment and sequential positioning within the accompanying speech (see the "segment" tier in Figure 8). At this stage, the coders assigned each 2HAS gesture a superordinate meaning category—delineated, holistic, or unclear. In addition, the more specific intentional objects were specified for both delineated and holistic meanings (see the "subcategory" tier in Figure 8). For the subsequent step of intersubjective corroboration, we also translated the transcribed speech into English. For data from the Swedish group, this entire

second step was carried out by the Swedish research partner. Figure 8 displays the ELAN screen containing all the information annotated during these two steps.

In the final third step, I and the research partner cross-checked all the assigned superordinate meaning categories and their corresponding intentional objects for the 2HAS gestures that had been annotated by the other coder in the second step. In cases of disagreement, the original coder clarified their initial interpretation, the other provided suggestions, and the two discussed the instance until reaching a consensus. The results presented in Chapter 4 are all those that have undergone this intersubjective validation step.

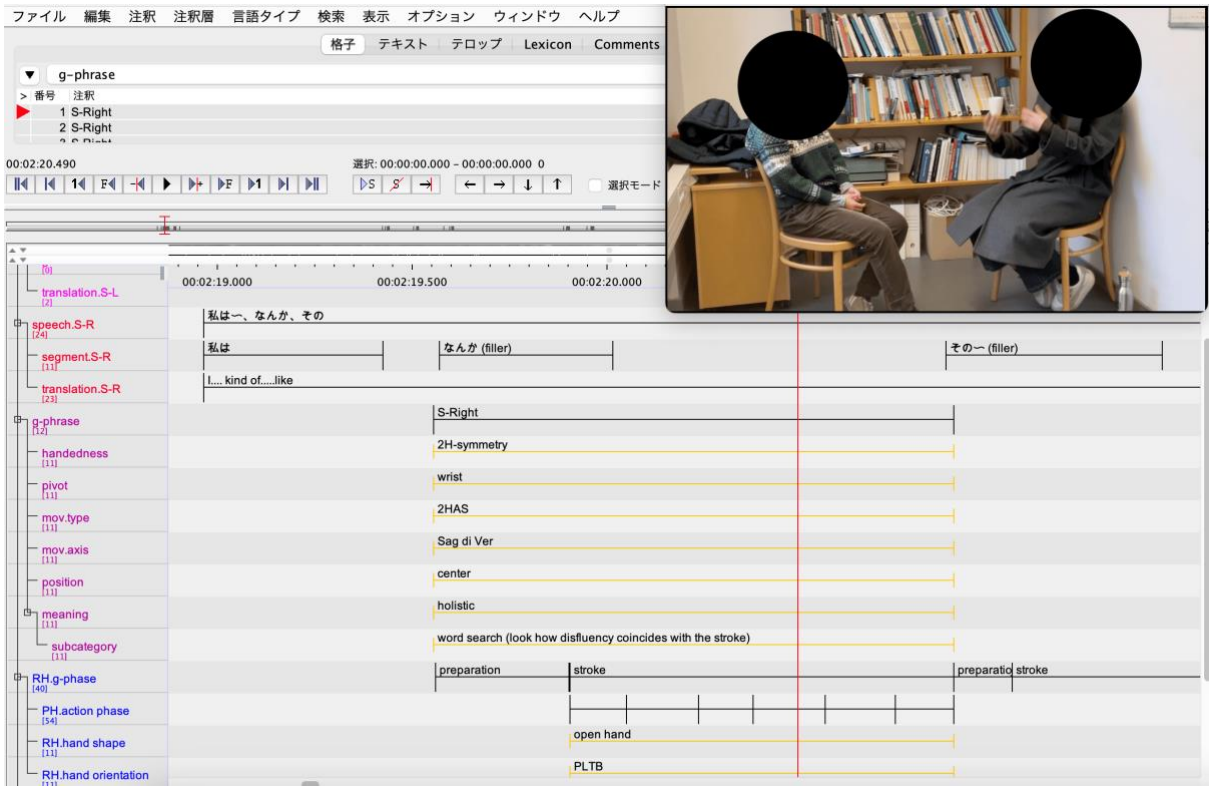


Figure 8. An ELAN screen containing all the information annotated during the first two steps

Chapter 4. Results and discussion

This chapter presents the results of the empirical analysis conducted on the 2HAS gesture, based on eight hours of dialogic data per cultural group—yielding a total of 16 hours of recorded conversation. Within this corpus, Japanese participants produced the 2HAS gesture 71 times, while their Swedish counterparts used it 101 times. The chapter is organized into three main sections. Section 4.1 examines the meanings of 2HAS, comparing points of convergence and divergence between the two language groups. Section 4.2 shifts focus to frequency—both across the two corpora and among individual speakers—exploring how distributional patterns may reflect the interplay of genetic and generative sedimentation. Section 4.3 then synthesizes these findings into a broader theoretical discussion in light of the Motivation & Sedimentation Model (MSM), returning to the questions concerning the nature of recurrent gestures, embodiment and sedimentation.

4.1. Meaning of 2HAS

This section examines the meaning of 2HAS gestures, thereby addressing RQ1. Figure 9 visualizes, for each cultural group, the proportion of 2HAS tokens that were coded as delineated, holistic, or unclear. In the Japanese data, 51 of 71 tokens (72 %) were found to convey delineated meanings, 16 tokens (23 %) expressed holistic meanings (e.g., stance marking or word-search; see Section 4.1.2), and 4 tokens (6 %) were classified as unclear. The Swedish distribution was similar in its broad outline—66 of 101 tokens (65 %) were delineated and 5 tokens (5 %) unclear—yet it showed a noticeably larger share of holistic meanings (30 tokens, 30 %).

Taken together, Figure 9 confirms that, in both communities, 2HAS was most often deployed to signify specific parts of a larger whole (e.g., proposition), but it can also be used to qualify, modulate, or express epistemic stance toward that whole. The following subsections investigate the nature of these delineated and holistic meanings in more detail.

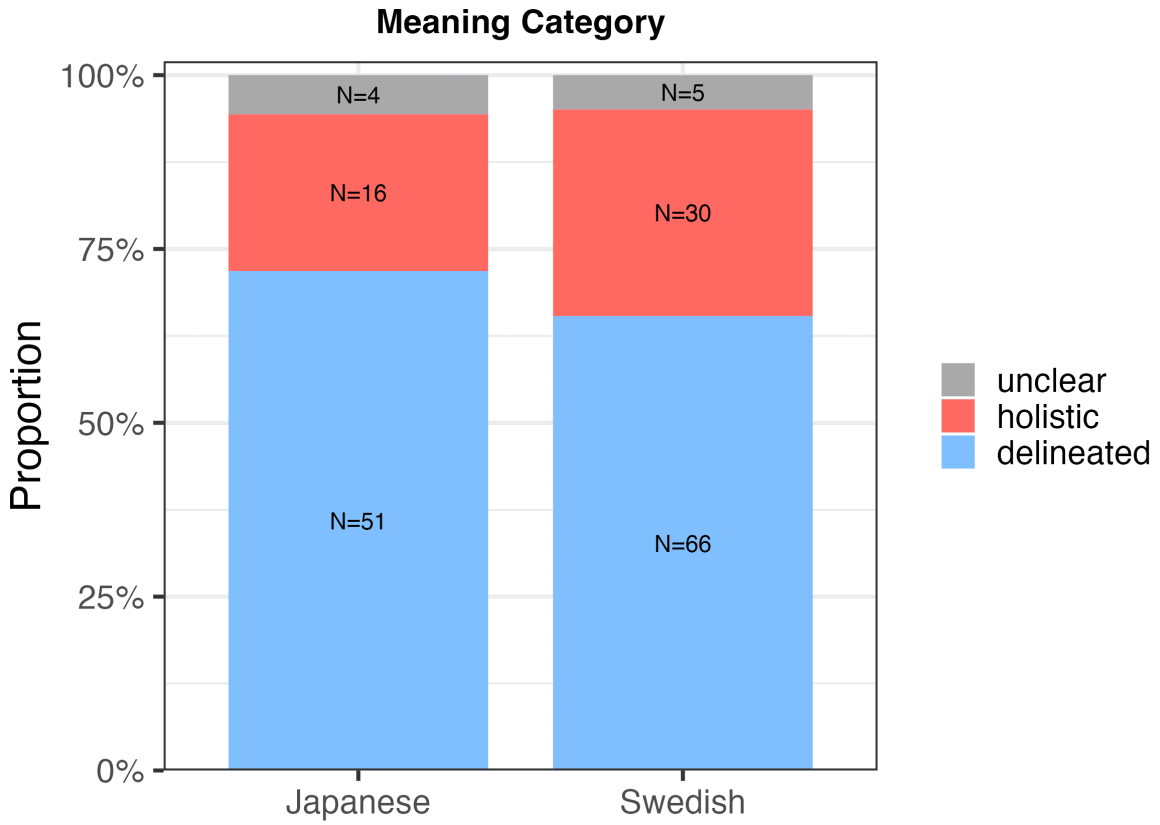


Figure 9. Distribution of delineated, holistic, and unclear meanings of 2HAS in Japanese and Swedish.

4.1.1. Delineated meaning of 2HAS

Figure 10 breaks down the delineated intentional objects signified by 2HAS into five schematic subcategories: BIDIRECTIONALITY, MULTIDIRECTIONALITY, GIVE & RECEIVE, (INCREMENTAL) PROCESS, and OTHER. This classification was carried out based not only on the content of co-occurring speech, but also on the formal features of the gesture itself—e.g., hand shape, hand orientation and gesture position—as discussed in Section 2.3.2. For example, a 2HAS gesture with extended index fingers typically emphasizes the BIDIRECTIONAL path of an action (e.g., communicating or moving back and forth), aligning with an image schematic interpretation. In contrast, open cupped hands often highlight the action itself, such as GIVE & RECEIVE, suggesting mimetic schemas grounded in instrumental action. Importantly, as emphasized by Zlatev and Blomberg (2016), these schematic delineated intentional objects signified by 2HAS are not the referents of these gestures but rather function as motivational structures of the gestural enactment.[20]

[20] For instance, if a 2HAS gesture co-occurs with speech such as *I traveled back and forth between*

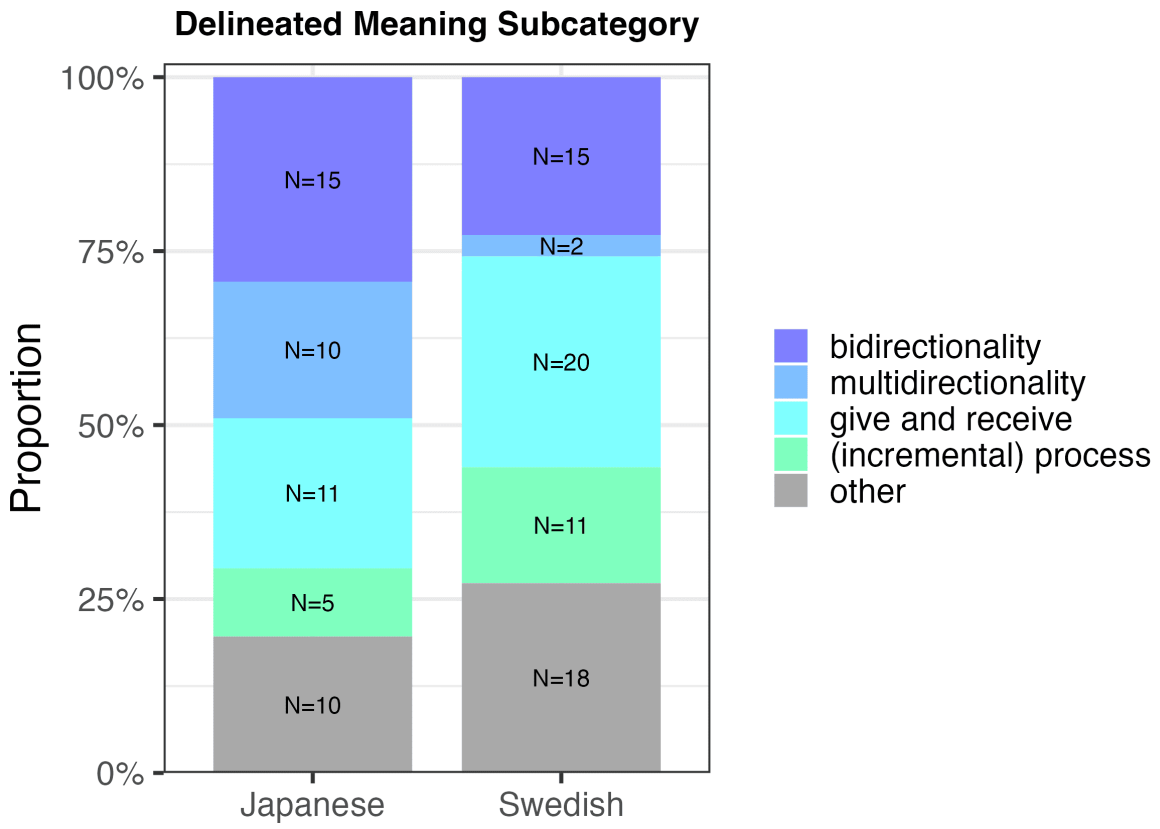


Figure 10. Distribution of subcategories of delineated meaning in Japanese and Swedish

As shown above, BIDIRECTIONALITY and GIVE & RECEIVE were the most frequently assigned subcategories in both groups (except for OTHER in Swedish). In Japanese, 15 instances (29%) were classified as BIDIRECTIONALITY, and 11 (22%) as GIVE & RECEIVE. Similarly, in Swedish, 15 instances (23%) fell under BIDIRECTIONALITY, and 20 (30%) under GIVE & RECEIVE. These two categories together account for more than half of all delineated meanings of 2HAS in both cultural groups.

In addition, MULTIDIRECTIONALITY, which was not introduced in the earlier chapters, covers instances in which 2HAS co-occurs with speech about going to various places, moving in diverging directions, or similarly dispersed trajectories. In such cases the gesture foregrounds *multiple* vectors rather than the two-pole reciprocity of BIDIRECTIONALITY, yet—because of its visuo-spatial nature—it can still be regarded as having an image-schematic structure, similar to BIDIRECTIONALITY. This subcategory was observed in 10 Japanese cases

Japan and China, the actual referent denoted by the whole utterance including the 2HAS gesture is the act of back-and-forth travel between the two countries, while the intentional object of the 2HAS gesture is BIDIRECTIONALITY.

(20%) but only 2 in the Swedish data (3%). Notably, when this subcategory appeared in the Japanese data, it frequently co-occurred with an adjective *ironna* ('various', 'many different'), as illustrated in Figure 11.

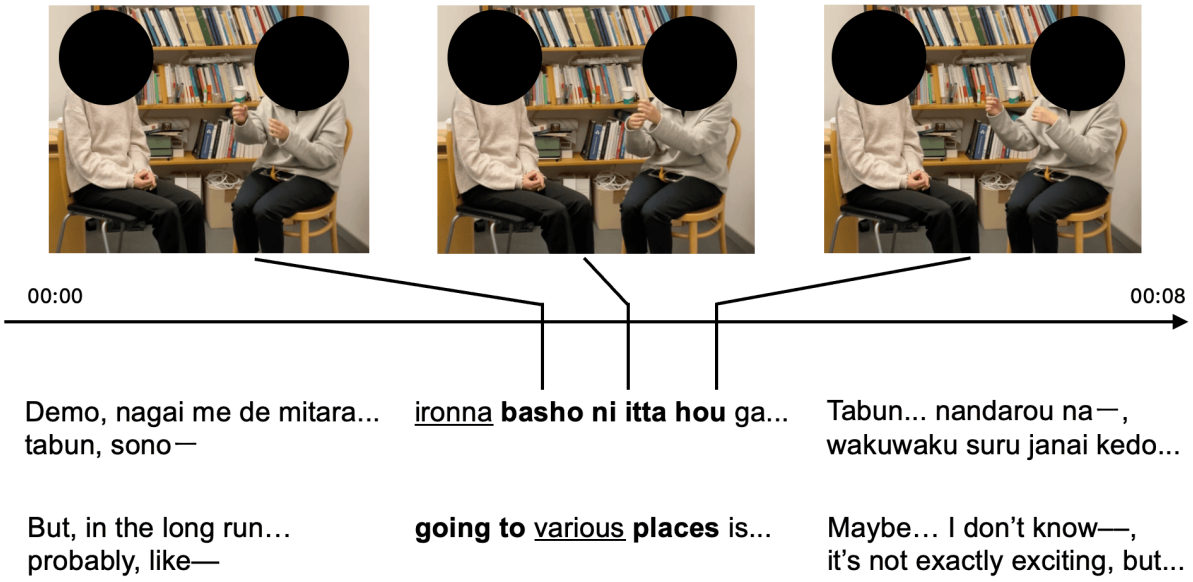


Figure 11. Example of a MULTIDIRECTIONALITY 2HAS gesture performed by a Japanese speaker. **Bold** indicates the stroke phase; underline marks the preparation phase.

Furthermore, the (INCREMENTAL) PROCESS category captures uses where speakers talk about gradual development or ongoing accumulation (e.g., *utvecklas intill* in Swedish, 'developing until'; *kaseide kaseide kaseide* in Japanese, 'earning and earning and earning', see Figure 12). Notably, a large proportion of these tokens were produced in tandem with a cyclic gesture, where the two hands alternate along a circular path. Across the whole dataset (including the holistic and unclear instances), a cyclic + 2HAS combination appeared 19 times in Japanese (27 % of all Japanese 2HAS tokens) and 21 times in Swedish (21 %). Within the (INCREMENTAL) PROCESS sub-set specifically, 3 of 5 Japanese tokens and 8 of 11 Swedish tokens were cyclic + 2HAS.

Interestingly, cyclic + 2HAS gestures also occurred in the OTHER category: 5 of 10 Japanese and 5 of 18 Swedish instances. By contrast, this combination was rare in the core categories of BIDIRECTIONALITY, MULTIDIRECTIONALITY, and GIVE & RECEIVE: only 5 of 36 Japanese and 1 of 37 Swedish tokens in these three categories included a cyclic component. This suggests that the sense of incremental unfolding owes more to the cyclic movement pattern itself—long associated with PROCESS in gesture research (Ladewig, 2011; McNeill, 1992)—than to the alternation along the sagittal axis per se.

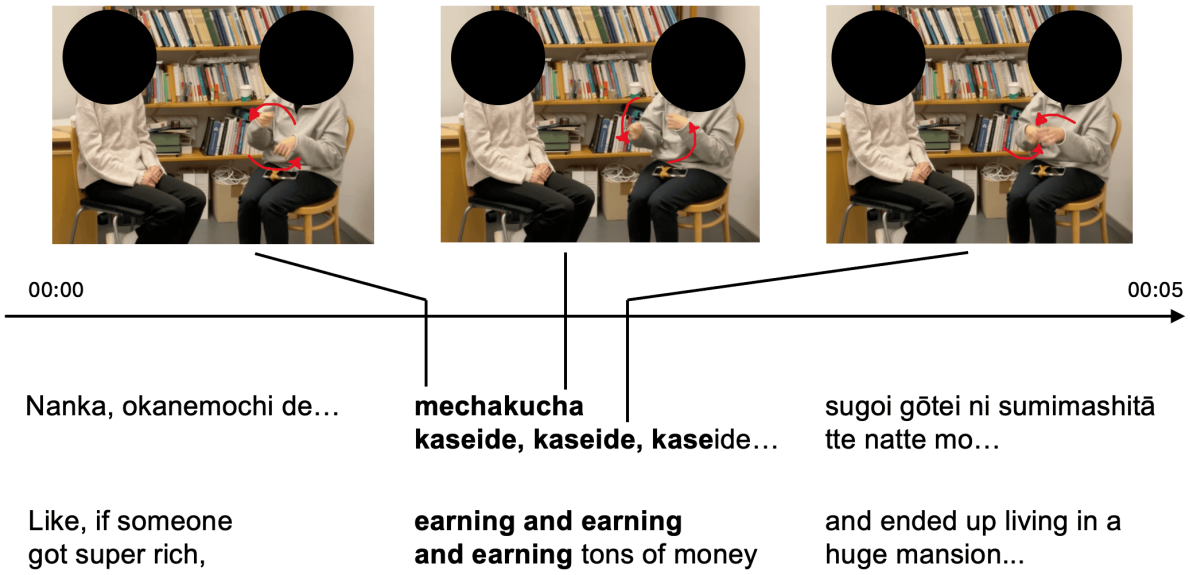


Figure 12. Example of a (INCREMENTAL) PROCESS cyclic + 2HAS gesture performed by the same Japanese speaker shown in Figure 11.

Taken together with MULTIDIRECTIONALITY, the subcategories that were a priori hypothesized as embodied motivations—BIDIRECTIONALITY and GIVE & RECEIVE—accounted for the majority of all delineated meaning tokens in both groups (71 % in Japanese; 56 % in Swedish). The scarcity of cyclic + 2HAS within these main categories reinforces the idea that the core of the delineated meanings of 2HAS is dominated by its pure form—not mixed with cyclic movement. This supports the hypothesis that the image and mimetic schemas of BI-/MULTI-DIRECTIONALITY and GIVE & RECEIVE were not only recurring but also *central* motivations for the use of 2HAS across both the two groups, and thus potentially across the cultures (to the extent that the findings of the study can be shown to be representative). This finding provides relatively strong evidence that, at least in its delineated meanings, 2HAS is grounded in pan-human aspects of embodiment and iconicity. That is, the form of the gesture directly re-enacts bodily experiences of moving back and forth or exchanging objects—basic interactions shared across human cultures. These results resonate with the predictions made in Section 2.4, where such cross-cultural commonality in delineated meanings was anticipated.

At the same time, however, the OTHER category accounted for a non-negligible portion of the data: 10 of 51 (20%) in Japanese and 18 of 66 (27%) in Swedish. These cases could not be clearly attributed to any of the four schematic subcategories and may thus reflect alternative motivating forces beyond those rooted in image or mimetic schemas. The relatively high proportion of these unclassified instances suggests that, while embodied schemas explain a

majority of delineated meanings of 2HAS, they do not account for *all* of them.

Finally, although the distribution of MULTIDIRECTIONALITY varied between groups (20% Japanese vs. 3% Swedish), the sample size is too small for a definitive cross-cultural conclusion. Nonetheless, this discrepancy might reflect different cultural tendencies in conceptualizing spatial multiplicity, a question that merits further investigation.

4.1.2. Holistic meaning of 2HAS

The present subsection turns from the delineated meanings to the ways in which 2HAS *qualifies* an utterance as a whole—its holistic meanings. Figure 13 visualizes how six inductively and intersubjectively derived subcategories—UNCERTAINTY, OPTIONALITY, WORD SEARCH, CONTRASTIVE, OBVIOUSNESS, and OTHER—are distributed across the Japanese and Swedish data.

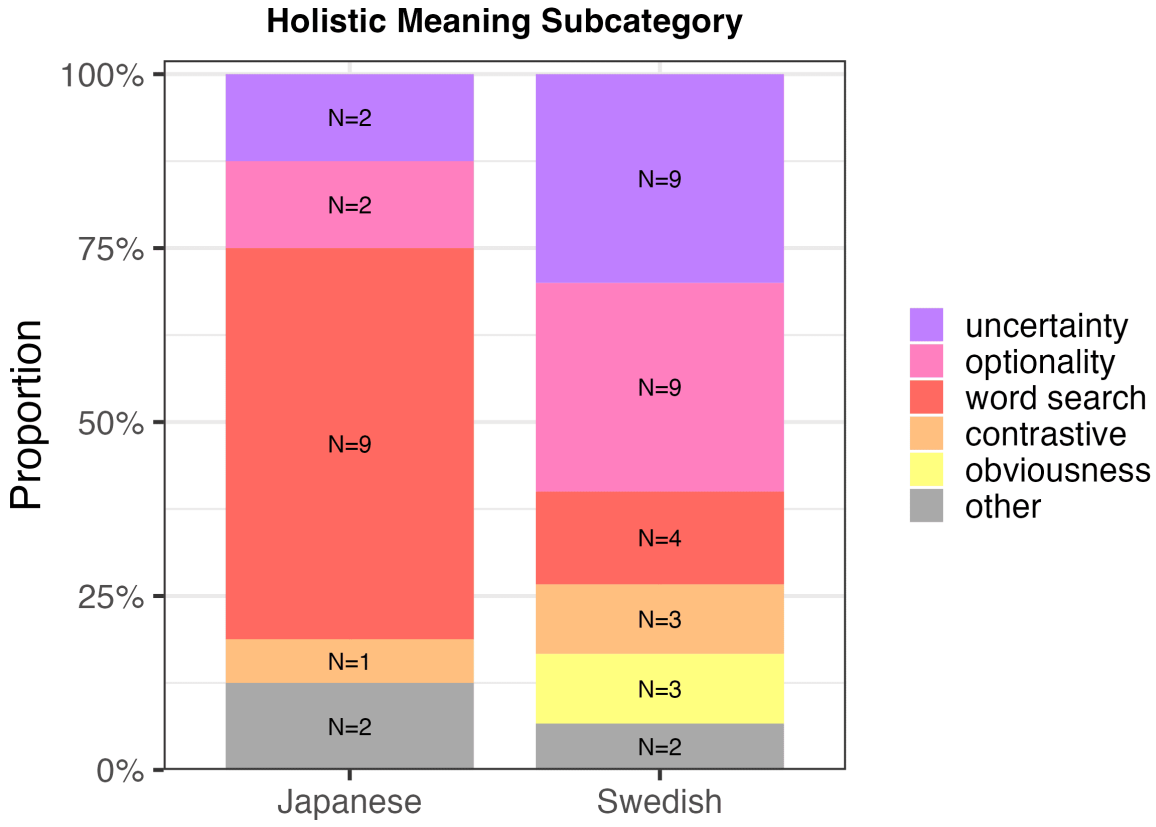


Figure 13. Distribution of subcategories of holistic meaning in Japanese and Swedish

Crucially, none of these labels was imposed a priori; instead, they emerged through an intersubjective coding dialogue with the research partner, in keeping with the

phenomenological triangulation procedure outlined in Chapter 3.[21] The paragraphs that follow define each category and interpret the cross-cultural patterning evident in Figure 13.

The UNCERTAINTY category (see Figure 14 for an example) signifies a speaker's epistemic stance of lacking confidence in the propositional content of their utterance. As anticipated in Section 2.4, this holistic meaning aligns with Bressem and Müller's (2014) characterization of 2HAS, which emphasized UNCERTAINTY as part of its semantic core—particularly in German, a language culturally and linguistically closer to Swedish. In the Swedish group, 2HAS was used to signify UNCERTAINTY in 9 out of 30 cases, comprising 30% of all Swedish tokens. In contrast, the Japanese group exhibited this category only twice, accounting for just 13% of their total. This disparity supports our theoretical prediction from Section 2.4: holistic meanings such as epistemic stance are more likely to reflect culturally sedimented conventions, and thus diverge across groups. Interestingly, in the Swedish data, 2HAS in this function frequently co-occurred with lexical hedges such as *liksom* ('like', 'kind of'), reinforcing its role as a polysemiotic epistemic hedge. This co-occurrence suggests functional overlap with linguistic expressions of UNCERTAINTY, further confirming the epistemic nature of this gesture in the Swedish context.

Figure 14. Example of an UNCERTAINTY 2HAS gesture performed by a Swedish speaker.

[21] After assigning a superordinate meaning category (delineated vs. holistic), each of us described how the 2HAS gesture qualified the utterance as a whole, using our native speaker intuitions (first-person). I then examined all the instances across the groups and looked for commonalities, which led to the formulation of subcategories. Referring to the defining criteria for each subcategory, the Swedish research partner subsequently reviewed all instances across the groups, thereby confirming intersubjective validity (second-person). Cases of disagreement were discussed until a consensus was reached.

The OPTIONALITY category (see Figure 15 for an example) signifies the speaker's presentation of possibilities, potential consequences, or hypothetical examples. In such cases, 2HAS was used to mark the content as non-exclusive, inviting the addressee to consider several options. This category accounted for 9 instances (30%) in the Swedish group and 2 (13%) in the Japanese group. Strikingly, 10 out of these 11 occurrences were performed using an open-hand configuration. This hand shape, combined with the sagittal alternation of 2HAS, gives the impression of presenting options or offering alternatives. Thus, the movement can be interpreted as enacting the *giving* or *showing* of possibilities, potentially rooted in mimetic schemas of these manual actions. While the OPTIONALITY category shares some surface similarities with UNCERTAINTY, it differs fundamentally in the speaker's stance. Whereas UNCERTAINTY involves a lack of confidence or commitment, speakers using 2HAS as an OPTIONALITY marker appear confident. They are not unsure about what they are saying; rather, they are deliberately opening up a semantic space of possibilities. This distinction highlights a key nuance in the functional spectrum of holistic meanings and also underscores the importance of embodied construal in shaping even holistic meanings of the 2HAS gesture.

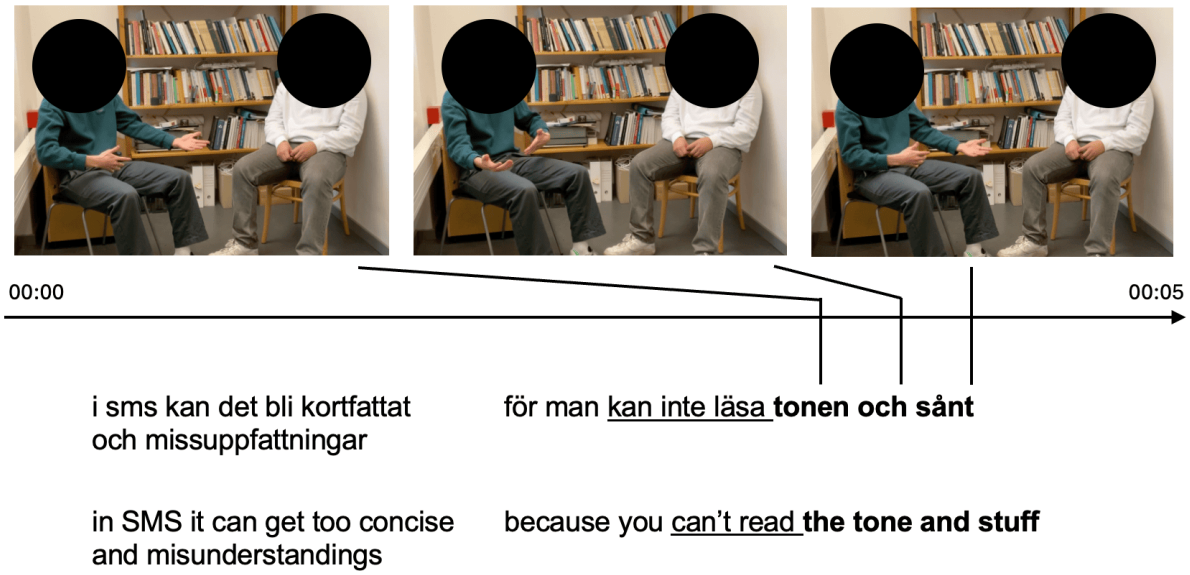


Figure 15. Example of an OPTIONALITY 2HAS gesture performed by a Swedish speaker.

As its name suggests, the WORD SEARCH category surfaces when the speaker is searching for the right expression or idea. Although the boundary between this category and UNCERTAINTY was not always clear-cut, specific criteria were employed during coding to maintain analytic clarity. These included: (i) temporal coincidence of 2HAS with speech

disfluency (see Figure 8); (ii) slowing of speech rate prior to the gesture; and (iii) co-occurrence with lexical phrases such as *nandakke* in Japanese and *vad heter det* in Swedish ('what's it called'). Empirically, this category was most prominent in the Japanese group, where 9 out of 16 cases (56%) were categorized as WORD SEARCH. In the Swedish group, this category appeared only 4 times, representing 13% of their total. Importantly, 5 of the 13 total WORD SEARCH gestures (3 in Japanese, 2 in Swedish) were mixed forms combining 2HAS with cyclic movements (cyclic + 2HAS), suggesting that this meaning may be partly rooted in the cyclic component in such cases—consistent with the findings that WORD SEARCH is a dominant use for cyclic gestures in German (Ladewig, 2011). Nonetheless, the remaining eight instances featured a "pure" 2HAS trajectory, indicating that the gesture schematically represents a mental process of moving "back and forth" through conceptual space—potentially an image-schematic motivation grounded in bodily experience and visual perception of BIDIRECTIONALITY.

In addition to the three major subcategories described above, CONTRASTIVE uses of 2HAS can be considered to fall under Kendon's (2004) discourse-parsing functions: they juxtapose two alternatives, often aligning with lexical phrases such as *on the other hand* or *either A or B*. In the data, this category appeared only once in the Japanese group and three times in the Swedish group. Despite its relative infrequency, its semiotic function is clear and distinct. The alternation in the sagittal axis naturally supports a contrastive reading, as it rhythmically parses discourse into segments, guiding the listener's attention between two options.

Another minor subcategory was OBVIOUSNESS, which appeared exclusively in the Swedish group (3 instances). In these cases, 2HAS signifies that the speaker regarded the content as self-evident or commonsensical. At first glance, this epistemic stance seems to contradict the UNCERTAINTY category, which was otherwise dominant in the Swedish data. However, closer inspection revealed that 2 out of these 3 instances were performed with the Palm Up Open Hand (PUOH; Müller, 2004) configuration—a gesture previously identified as an OBVIOUSNESS marker in earlier studies (see Marrese et al., 2021, for an extensive discussion). Therefore, it is likely that the OBVIOUSNESS meaning may not be rooted in 2HAS per se, but rather in the PUOH component.

Finally, each group contained 2 instances of 2HAS that did not fit into the five categories above. These were classified as OTHER, encompassing idiosyncratic or ambiguous cases.

From the overall pattern shown in Figure 13, two complementary interpretations are possible, each bearing on the hypothesis of culturally contingent divergence presented in Section 2.4. The first interpretation supports the hypothesis that holistic meanings of 2HAS are

culturally specific and reflect generative sedimentation. Japanese speakers predominantly used 2HAS for WORD SEARCH (56%), while Swedish speakers favored UNCERTAINTY and OPTIONALITY (30% each), jointly comprising 60% of their cases. These divergent patterns suggest that certain holistic meanings have become sedimented differently across speech-gesture communities. Specifically, this divergence may be influenced by competition from other recurrent gestures. As discussed above, cyclic gestures dominate the WORD SEARCH function in German (Ladewig, 2014), potentially limiting the space for 2HAS in this domain in Swedish as well. In Japanese, however, the potential lack of such a strong competitor may have enabled 2HAS to fill this communicative niche. These findings indicate that 2HAS's holistic meanings are shaped by local histories of generative sedimentation and interactional norms.

The second interpretive perspective presents evidence against the cultural divergence hypothesis. With the exception of the small OBVIOUSNESS subset, all subcategories of holistic meaning—UNCERTAINTY, OPTIONALITY, WORD SEARCH and CONTRASTIVE—appeared in both groups. Even though the proportions differ, their existence across both cultural settings suggests that these meanings are not entirely culture-specific. Rather, they may draw upon embodied motivations shared across human communicative practices. As discussed above, WORD SEARCH gestures may reflect the image-schematic structure of conceptual back-and-forth movement, while OPTIONALITY markers may embody the mimetic schema of GIVE or SHOW. Even UNCERTAINTY, though less obviously grounded in action schemas, may be visualized as a kind of cognitive oscillation—a conceptual motion in multiple directions.

Taken together, these findings align well with Cooperrider's (2019) characterization of recurrent gestures as "natural conventions"—gestures that are culturally selected, in different degrees of commonality, from a menu of embodied, motivated possibilities. This dual nature—embodied yet conventionalized—helps explain both the similarities and differences in holistic meaning of 2HAS across cultural groups.

Combined with the findings from Section 4.1.1 on delineated meanings, we now have a more comprehensive picture of how 2HAS operates across the two distinct cultural groups. The results indicate that while delineated meanings tend to exhibit cross-cultural convergence, holistic meanings reveal at least partially culture-specific sedimentations. Yet both types of meaning are rooted, albeit to different degrees, in embodied experience. This dialectical interplay between embodiment and sedimentation is further explored in Section 4.3, where we will revisit the theoretical models of gestural meaning-making in light of the Motivation &

Sedimentation Model (MSM) introduced in Section 2.1.4.

4.2. Frequency and individual differences of 2HAS

The preceding section disentangled the delineated and holistic meanings of the 2HAS gesture. We can now step back from micro-analysis to ask how often 2HAS surfaces in discourse and how evenly it is distributed across individual speakers in the two speech-gesture communities investigated. Frequency patterns are not mere background statistics: they provide indirect evidence for the degree to which a practice has become sedimented in a community and for the extent to which it is internalized as an individual habit. This section therefore combines descriptive statistics with a Bayesian regression analysis to illuminate both *generative sedimentation* (community-wide conventionalization) and *genetic sedimentation* (individual habituation) of the 2HAS gesture.

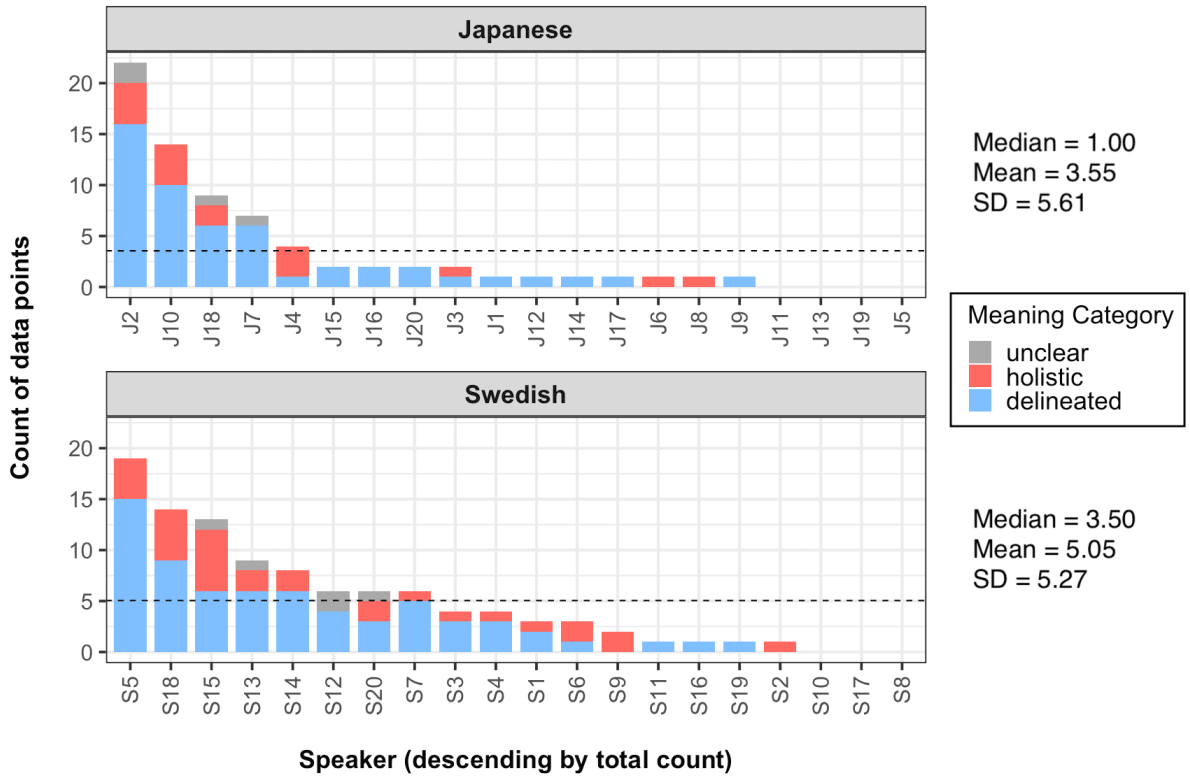


Figure 16. Distribution of 2HAS gesture tokens by speaker and meaning category in Japanese and Swedish

The bar plot presented in Figure 16 provides a comprehensive overview of how 2HAS gestures were distributed among individual speakers within each cultural group. In the upper

panel, Japanese participants (J1–J20) are displayed in descending order of total 2HAS occurrences. The lower panel shows the same for Swedish speakers (S1–S20). The black dashed line in each panel indicates the mean count per group. Quantitatively, Japanese speakers show a strong right-skewed distribution, with the majority contributing very few data points and a small minority accounting for most of the observed 2HAS occurrences. Specifically, the Japanese group shows a mean of 3.55, a median of 1.00, and a standard deviation (SD) of 5.61, indicating high dispersion and strong individual differences. The Swedish group also displays a right-skewed distribution but with a less extreme long tail: the mean is 5.05, the median is 3.50, and the SD is 5.27.

Among the Swedish speakers, eight individuals exceed a total count of 5, and twelve exceed a total of 3. In contrast, only four Japanese speakers exceed 5, and five exceed 3. This suggests that more Swedish speakers consistently incorporate 2HAS into their gestural repertoire. Furthermore, since highly conventionalized emblems such as THUMBS UP or AIR QUOTES did not appear even once in this dataset (as far as observed), it would be misleading to equate low frequency with low degree of conventionalization. In fact, given that even a single occurrence of 2HAS might suggest its habitual or genetically sedimented status within the individual's communicative repertoire, instances with a count as low as one or two should not be dismissed. With this in mind, the finding that 16 Japanese speakers and 17 Swedish speakers each used 2HAS at least once is highly informative. This suggests that 2HAS has undergone some degree of generative sedimentation in both communities—having been socially transmitted and stabilized to the extent that it appears in the repertoire of the majority of participants. However, the higher average and median counts among Swedish participants imply a stronger degree of generative sedimentation.

To statistically assess such group-level differences, a Bayesian negative-binomial regression (Winter & Bürkner, 2021) was fitted using the brms package (Bürkner, 2017) in R, with *total count per speaker* as the response variable and *group* (Japanese = 0, Swedish = 1) as the predictor. The choice of a Bayesian framework is appropriate due to the small sample size (N = 40) and the inferential flexibility it offers (Georgiou, 2024). Unlike frequentist methods, which often compel researchers to make a dichotomous decision between "significant" and "non-significant," Bayesian analysis allows us to directly assess the likelihood of specific parameter values given the data (Levshina, 2016). Furthermore, analogous to the conceptual-empirical loop, Bayesian modelling allows for "the incorporation of prior knowledge via the specification of prior distributions, which has several advantages for more theory-guided

statistical modelling" (Winter & Bürkner, 2021, p. 6). The model used in the present thesis employed the following weakly informative priors, since using weakly informative priors is generally recommended, particularly when dealing with small samples (Winter & Bürkner, 2021):

b ~ Normal(0, 0.5); Intercept ~ Normal(0, 2.5); shape ~ exponential(1).

Four chains were run with 4,000 iterations each, including 2,000 warm-up iterations. A seed (1234) was set for reproducibility. Figure 17 plots the posterior distribution of the group coefficient.

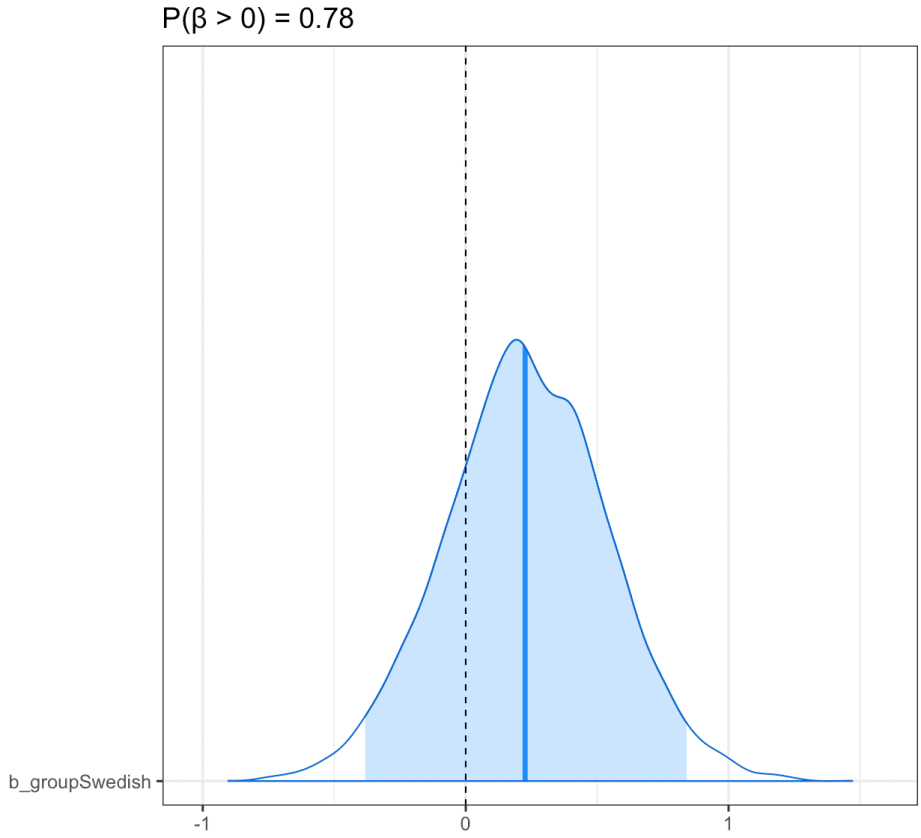


Figure 17. Posterior distribution of group effect (Japanese vs. Swedish) on 2HAS count

The intercept (corresponding to the Japanese group) is 1.34. This intercept indicates that Japanese speakers are expected to produce exp(1.34) ≈ 3.82 tokens of 2HAS per speaker. The estimate for the Swedish group coefficient is 0.23 (95% CrI = [–0.38, 0.84]), corresponding to a rate ratio of exp(0.23) ≈ 1.26. Thus, Swedish speakers are expected to produce approximately 1.26 times as many 2HAS tokens as Japanese speakers. The posterior probability that this effect

is positive is $P(\beta > 0) = 0.78$. Therefore, although the 95% credible interval crosses zero, the posterior probability of 0.78 suggests *a modest tendency for Swedish speakers to perform 2HAS more frequently than Japanese speakers*,[22] which is in line with the interpretation of descriptive statistics presented in Figure 16.

This inference—that generative sedimentation is more pronounced in the Swedish group—aligns well with the results in Section 4.1.2, where holistic meanings of 2HAS were found to be more prevalent among Swedish speakers. Indeed, nearly twice as many Swedish participants (N = 13) used 2HAS with holistic meanings compared to their Japanese counterparts (N = 7). This points to a greater degree of cultural diffusion and community-wide adoption of the 2HAS gesture's more extended, non-literal meanings within the Swedish group.

While the preceding analysis has focused primarily on generative sedimentation, it is essential also to consider *genetic sedimentation*—habit formation based on individual embodied experience. The high standard deviations observed in both groups (Japanese: 5.61, Swedish: 5.27) point to substantial individual variability. Such variation strongly suggests that 2HAS usage is not solely a matter of cultural normativity but is also deeply shaped by embodied dispositions and individual communicative habits. This supports the earlier claim in Section 2.1.4, where I revised the MSM framework to argue that "habit-formation in an individual body always already takes place in a community saturated with Sedimented norms." Thus, 2HAS is best characterized as a *genetically sedimented practice conditioned by generative sedimentation*—an embodied gesture whose recurrence is shaped by both social norms and individual habits.

4.3. Recurrent gestures, embodiment and sedimentation (back again)

This section synthesizes the empirical findings and theoretical discussions presented in Sections 4.1 and 4.2. Drawing upon the revised version of the MSM introduced in Section 2.1.4, this section offers a comprehensive conceptual discussion on how embodiment and sedimentation interplay within the recurrent gestural practice of 2HAS, thus returning to the broader theoretical concerns of this thesis through the lens of the conceptual-empirical loop.

In response to Blomberg's (2019, p. 81) question regarding "whether these [i.e., genetic and generative sedimentation] can be subsumed under one concept or should be treated as two

[22] However, given the level of uncertainty, this result should be interpreted with caution.

distinct concepts," the detailed analysis of delineated and holistic meanings of 2HAS presented in Section 4.1 demonstrates that distinguishing between the two is not only theoretically necessary but empirically grounded. More specifically, with respect to their relevance to embodiment—manifested in terms of the degree to which the motivational origin is forgotten—the empirical semiotic analysis in Section 4.1 confirmed prior theoretical considerations. Namely, genetic sedimentation remains closely connected to bodily experience—thereby retaining a link to embodiment—whereas generative sedimentation involves the replacement of such bodily foundations by conventionality and symbolicity.

This distinction can be further elaborated through the concepts of *primary iconicity* and *secondary iconicity*. According to Sonesson (2010, p. 39, his emphasis),

> A primary iconic sign is a sign in the case of which *the perception of a similarity* between an expression E and a content C *is at least a partial reason* for E being taken to be the expression of a sign the content of which is C. That is, iconicity is really the motivation (the ground), or rather, one of the motivations, for positing the sign function. A secondary iconic sign, on the other hand, is a sign in the case of which *our knowledge that E is the expression of a sign* the content of which is C, in some particular system of interpretation, *is at least a partial reason* for perceiving the similarity of E and C. Here, then, it is the sign relation that partially motivates the relationship of iconicity.

In the context of 2HAS, if we interpret Sonesson's notion of C (= content) as encompassing both the delineated and holistic intentional objects signified by 2HAS, then this distinction between primary and secondary iconicity maps neatly onto the contrast between the delineated and holistic meanings analyzed in Section 4.1.

For delineated meanings, the motivation is transparent and predictive. This is precisely why it was possible to hypothesize and then empirically identify specific embodied schemas—such as the BIDIRECTIONALITY image schema and the GIVE & RECEIVE mimetic schema—as subcategories of delineated meaning before beginning the actual analysis. In Sonesson's terms, the similarity between the gestures (E) and these embodied schemas (C) serves as the basis for the E-C sign relation. Importantly, these embodied schemas emerge from individuals' bodily engagement with the world and others—that is, from genetic sedimentation—and are publicly available through embodied intersubjectivity (see also Section 2.1.2).

In contrast, for holistic meanings, such transparent motivations are difficult to anticipate a

priori. Instead, their subcategories were generalized through an inductive process, and it was only through retrospective analysis that the potential iconic bodily motivations—prior to their generative sedimentation—could be identified. Thus, the iconicity involved in holistic meanings of 2HAS is secondary in the sense that symbolicity is primary here: it is symbolicity that enables the perception of similarity between E and C. This secondary iconicity—or alternatively, primary symbolicity—resonates with one interpretation of the findings presented in Section 4.1.2, namely that the holistic meanings of 2HAS are primarily shaped by local histories of generative sedimentation.

Building upon this insight, Section 4.2 expanded the discussion by analyzing individual- and group-level frequency differences in the use of 2HAS. The results indicated that although 2HAS displays aspects of conventionalization (generative sedimentation) across groups—particularly among Swedish speakers—there is still substantial inter-individual variation in its actual usage. This reinforces the idea that while 2HAS is shaped by culturally sedimented norms, its practice remains fundamentally grounded in bodily experience. As such, 2HAS should be understood as a primarily genetically sedimented practice, conditioned—but not fully determined—by generative sedimentation. This conclusion motivates a final reconsideration of the tentative placement of recurrent gestures within the MSM proposed in Section 2.1.4.

Given these findings, we can now reaffirm and refine the positioning of recurrent gestures—such as 2HAS—within the Embodied/Sedimented level of the revised MSM. Recurrent gestures are neither as purely spontaneous as gesticulations nor as fully conventionalized as emblems. Instead, they emerge from a *double motivation*—drawing from both pan-human bodily processes (at the Embodied level) and generatively sedimented conventions (at the Sedimented level). Over time, these dual sources become genetically sedimented within the individual, forming stable patterns of usage that still retain traces of their embodied origins. In this sense, recurrent gestures are best understood as structures of *embodied intersubjectivity* (e.g., Zlatev, 2017; Zlatev & Blomberg, 2016) situated at the Embodied/Sedimented level: stabilized enough to be recognizable across individuals and cultures, yet still flexible, personal, and bodily grounded. This conclusion is also visually supported by the conceptual diagram provided in Figure 18. As the arrows in the model indicate, recurrent gestures result from upward motivational forces from both the pure Embodied and Sedimented levels and are stabilized through genetic sedimentation. They thus constitute an intermediate layer in the human meaning-making—serving as a bridge between pre-signitive

bodily processes and symbolic social norms.

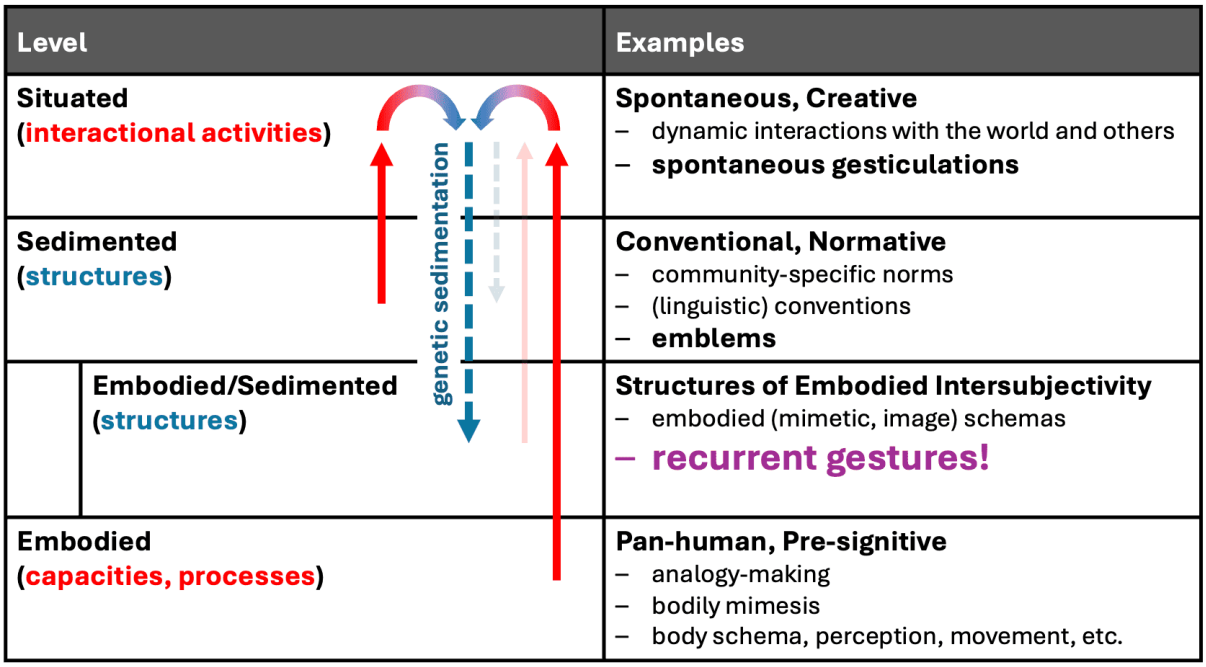


Figure 18. Theoretical placement of recurrent gestures within the revised MSM revisited

In summary, this chapter has demonstrated that 2HAS is a clear case of a recurrent gesture grounded in both embodiment and sedimentation. Through the conceptual-empirical loop, it has become evident that the distinction between genetic and generative sedimentation is not only theoretically sound but also empirically indispensable. The delineated meanings of 2HAS are shaped by genetic sedimentation and manifest primary iconicity, whereas holistic meanings are primarily shaped by generative sedimentation and exhibit secondary iconicity. Furthermore, the observed individual differences in gesture usage highlight the importance of recognizing recurrent gestures as genetically sedimented practices conditioned by generatively sedimented norms. These insights confirm the appropriateness of locating recurrent gestures within the Embodied/Sedimented level of the revised MSM, providing a robust conceptual framework for understanding the complex dynamics of recurrence in our gestural communication.

Chapter 5. Conclusion

This thesis set out to answer two closely linked questions formulated in Chapter 1:

RQ1: What meaning variations does 2HAS exhibit, particularly from a cross-cultural perspective?

RQ2: To what extent is 2HAS motivated by embodiment and to what by sedimentation, as reflected in its commonalities and differences across the cultural groups, respectively?

Building on the detailed analyses in Chapter 4, the present chapter summarizes and concludes the arguments presented in this thesis.

Regarding RQ1, across the 16-hour corpus, Japanese participants produced 71 tokens of 2HAS and Swedish participants 101. In both groups, a large majority of tokens served delineated meanings (Japanese = 72%, Swedish = 65%), confirming that speakers in both cultures most often performed the gesture to signify a specific part within a larger whole (e.g., a proposition). Within this layer, two subcategories—BIDIRECTIONALITY and GIVE & RECEIVE—accounted for over half of all delineated tokens in each group, indicating a robust common core grounded in pan-human experiences of reciprocal motion and manual exchange. Variation became visible only at the margins. Swedish speakers showed a slight preference for the GIVE & RECEIVE reading, whereas Japanese speakers produced proportionally more MULTIDIRECTIONALITY cases; yet these differences were minor and unlikely to affect mutual intelligibility. In short, 2HAS exhibits a cross-cultural semantic nucleus that is iconic, easily recoverable, and therefore highly convergent across the two communities investigated.

By contrast, the holistic meanings revealed the gesture's culturally nuanced edges. Four subcategories of holistic meaning were shared—UNCERTAINTY, OPTIONALITY, WORD-SEARCH, and CONTRASTIVE—yet their distribution diverged. Swedish speakers most often used 2HAS to hedge epistemic commitment (UNCERTAINTY) or invite alternatives (OPTIONALITY), whereas Japanese speakers preferred it as a visual expression of online lexical retrieval (WORD SEARCH). This dual-layered profile—a stable core of delineated meanings alongside a variable periphery of holistic meanings—answers RQ1: the meaning of 2HAS is both universal and shaped by local communicative conventions.

This makes a smooth transition to RQ2. On the one hand, the cross-cultural convergence in delineated meanings suggests strong motivation from pan-human embodied structures,

which are formed through individuals' bodily engagement with the world and others, i.e., *genetic sedimentation*, and are publicly available through embodied intersubjectivity. On the other hand, the divergent distribution of subcategories of holistic meaning between groups indicates a motivation rooted in *generative sedimentation*: repeated social use has conventionalized the gesture for group-specific discourse management. At the same time, the shared presence—albeit differently weighted—of each holistic subcategory supports Cooperrider's (2019) view of recurrent gestures as "natural conventions," selected from a pan-human repertoire of embodied possibilities. Put differently, embodiment sets the range; sedimentation shapes the local preferences, implying that gestural (as well as verbal) conventions are *not arbitrary*. Frequency data complemented this conclusion. Although the Swedish corpus contained more tokens overall, statistical analysis showed wide speaker-to-speaker variability in both groups, indicating that individual habit (genetic sedimentation) interacts with community norms (generative sedimentation) rather than being overridden by them.

In sum, the findings suggest a layered motivation: (1) a pan-human embodied base that secures cross-cultural stability for delineated meanings, and (2) culture-specific sedimented patterns that give rise to divergent holistic meanings. In the context of the revised MSM, this can be paraphrased as a *dual motivation* stemming from the pure Embodied and Sedimented levels, which is then gradually stabilized through the process of genetic sedimentation into a structure of embodied intersubjectivity at the Embodied/Sedimented level (see Figure 18 in Section 4.3). Through the comparative study of recurrent gestures, this thesis contributed to the empirical grounding of the theoretical revision of MSM by distinguishing between embodiment as pan-human capacities (the Embodied level), culture-specific conventions as products of generative sedimentation (the Sedimented level), and structures of embodied intersubjectivity as outcomes of genetic sedimentation (the hybrid Embodied/Sedimented level).

However, these conclusions—especially the observed commonalities across groups—should be interpreted in light of the fact that, although the Japanese participants knew almost no Swedish, they had lived in Sweden for more than six months and had been exposed to its culture; moreover, all of them expressed interest in contexts outside Japan and possessed at least B1 proficiency in English on the CEFR scale. These factors may have facilitated an implicit cultural convergence in gestural practices, potentially diminishing the otherwise observable cross-cultural contrasts. In other words, the shared delineated (and holistic)

meanings of 2HAS may partially reflect this group's embodied intersubjectivity being reshaped through extended intercultural experience. Therefore, while the embodied motivation for delineated meanings remains primary, the sedimentation of usage norms may also operate transnationally—particularly in increasingly mobile and globalized speech communities.

Even with this limitation, the present thesis makes a timely contribution to the study of recurrent gesture's universality and culture-specificity, a line of inquiry strongly encouraged but hitherto underdeveloped (Ladewig, 2024). Beyond gesture studies, the results speak to the larger cognitive-semiotic issue of how body, culture and sign co-operate: pan-human bodily experience supplies the iconic ground, social interaction sediments this ground into reusable patterns along with local communicative norms, and speakers draw on both layers—often simultaneously—to make meaning in real time.

In closing, I encourage a continued application of the conceptual-empirical loop of cognitive semiotics that has guided the inquiry into 2HAS throughout this thesis, thus promoting new theoretical insights, as well as empirical findings. The present findings demonstrate that even highly routinized signs such as recurrent gestures retain their corporeal origins, and that culture shapes meaning not by overwriting embodiment but by selectively amplifying its affordances. This conceptual refinement invites further empirical research into the dynamic interplay between embodiment and sedimentation in human semiosis.

References

Andrén, M. (2010). *Children's gestures from 18 to 30 months*. [Doctoral Thesis] Lund University.

Andrén, M. (2014). On the lower limit of gesture. In M. Seyfeddinipur & M. Gullberg (Eds.), *From gesture in conversation to visible action as utterance: Essays in honor of Adam Kendon* (pp. 153–174). Amsterdam: John Benjamins.

Blomberg, J. (2019). Interpreting the concept of sedimentation in Husserl's Origin of Geometry. *Public Journal of Semiotics, 9*(1), 78–94.

Blomberg, J. & Zlatev, J. (2014). Actual and non-actual motion: why experientialist semantics needs phenomenology (and vice versa). *Phenomenology and the Cognitive Sciences 13*(3), 395–418.

Brandt, Per A. (2011). What is cognitive semiotics? A new paradigm in the study of meaning. *Signata, 11*, 49-60.

Bressem, J. (2013), A Linguistic Perspective on the Notation of Form Features in Gestures, In C. Müller A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & S. Tessendorf (Eds), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction, vol. 1*, (pp. 1079–98), Berlin: de Gruyter Mouton.

Bressem, J., & Müller, C. (2014a). The family of AWAY gestures. Negation, refusal and negative assessment. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & J. Bressem (Eds.), *Body, Language, Communication. An international handbook on multimodality in human interaction (38.2)* (pp. 1592–1605). Berlin: De Gruyter Mouton

Bressem, J., & Müller, C. (2014b). A repertoire of German recurrent gestures with pragmatic functions. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill & J. Bressem (Eds.), *Body, language, communication: An international handbook on multimodality in human interaction (38.2)* (pp. 1575–1591). Berlin: De Gruyter Mouton.

Bressem, J., & Müller, C. (2017). The "Negative-Assessment-Construction" – A multimodal pattern based on a recurrent gesture? *Linguistics Vanguard, 3*(s1).

Bressem, J., & Wegener, C. (2021). Handling talk: A cross-linguistic perspective on discursive functions of gestures in German and Savosavo. *Gesture, 20*(2), 219–253.

Brookes, H., & Le Guen, O. (2019). Gesture studies and anthropological perspectives: An introduction. *Gesture, 18*(2–3), 119–141.

Bürkner, P. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28.

Cienki, A. (2013). Image schemas and mimetic schemas in cognitive linguistics and gesture studies. *Review of Cognitive Linguistics, 11*(2), 417–432.

Cienki, A. (2015). Spoken language usage events. *Language and Cognition 7*(4), 499–514.

Cooperrider, K. (2019). Universals and diversity in gesture: Research past, present, and future. *Gesture, 18*(2–3), 209–238.

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

Donald, M. (1991), *Origins of the Modern Mind: Three Stages in the Evolution of Human Culture*, Cambridge: Harvard University Press.

ELAN (Version 6.4) [Computer software]. (2022). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan.

Feyereisen, P., Van de Wiele, M., & Dubois, F. (1988). The meaning of gestures: What can be understood without speech? *European Journal of Cognitive Psychology, 8*, 3–25.

Fillmore, C. (1982). Frame Semantics. In The Linguistic Society of Korea (Ed.), *Linguistics in the Morning Calm*, (pp.111–137). Seoul: Hanshin.

Frege, G. (1948). Sense and Reference. *The Philosophical Review, 57*(3), 209–230.

Gallagher, S. (2005). *How the Body Shapes the Mind*, Oxford: Oxford University Press.

Gallagher, S. & Zahavi, D. (2008). *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science.* Routledge.

Georgiou, G. P. (2024). Bayesian Models Are More Sensitive than Frequentist Models in Identifying Differences in Small Datasets Comprising Phonetic Data. *Stats, 7*(4), 1483-1495.

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure.* Chicago: The University of Chicago Press.

Harrison, S. (2014). The organisation of kinesic ensembles related to negation. *Gesture, 14*(2), 117–141.

Harrison, S. & Ladewig, S. (2021). Recurrent gestures throughout bodies, languages, and cultural practices. *Gesture 20*(2). 153–179.

Harrison, S., Ladewig, S. H., & Bressem, J. (2021). The diversity of recurrency: A special issue on recurrent gestures. *Gesture, 20*(2), 143–152.

Hilpert, M. (2019). *Construction grammar and its application to English* (2nd ed.). Edinburgh: Edinburgh University Press.

Hinnell, J. (2018). The multimodal marking of aspect: The case of five periphrastic auxiliary constructions in North American English. *Cognitive Linguistics 29*(4), 773–806.

Højgaard Hansen, B. (2024). *Can machines (do) language? A cognitive semiotic exploration of Large Language Model-based systems, user practices and conceptions*. [Master's thesis] Lund University.

Jakobson, R. (1965), Quest for the Essence of Language, *Diogenes, 13*, 21–38.

Johnson, M. (1987). *The body in the mind: The bodily basis of reason and imagination*. Chicago: Chicago University Press.

Kendon, A. (1995). Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics, 23*(3), 247–279.

Kendon, A. (2004). *Gesture. Visible action as utterance*. Cambridge, UK: Cambridge University Press.

Kendon, A. (2017). Pragmatic functions of gestures: Some observations on the history of their study and their nature. *Gesture, 16*(2), 157–175.

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human computer interaction: International gesture workshop, Bielefeld, Germany, Sept. 17–19, 1997; Proc* (pp. 23–35). Berlin: Springer.

Kuryu, D. (2024a, August). *Flexible regularity and image-schematicity in multimodal constructions: The case of back and forth.* [conference presentation] The 13th International Conference on Construction Grammar (ICCG13). University of Gothenburg, Sweden.

Kuryu, D. (2024b, August). *Gestural homonymy leads to a polysemous multimodal construction: An account based on the Motivation & Sedimentation Model (MSM)*. [conference presentation] The fifth Conference of the International Association for Cognitive Semiotics (IACS-5). Lund University, Sweden.

Kuryu, D. (2025). Crossmodal collostructional analysis of English [ADV and ADV] constructions: Multimodal constructions or crossmodal collostructions? *Language and Cognition, 17*, e39. 1–28.

Ladewig, S. (2011). Putting the cyclic gesture on a cognitive basis. *CogniTextes, 6*. 1–26

Ladewig, S. (2014). Recurrent gestures. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill & J. Bressem (Eds.), *Body, language, communication: An international handbook on multimodality in human interaction (38.2)* (pp. 1558–1574). Berlin &

Boston: De Gruyter Mouton.

Ladewig, S. (2024). Recurrent gestures: Cultural, individual, and linguistic dimensions of meaning-making. In A. Cienki (Ed.), *The Cambridge Handbook of Gesture Studies*, (pp. 32–55). Cambridge: Cambridge University Press.

Ladewig S. (2025) Embodied sharpness: exploring the slicing gesture in political talk shows. *Frontiers in Psychology. 15*:1494192.

Lakoff, G. and M. Johnson (1980), *Metaphors We Live by*, Chicago: University of Chicago Press.

Lakoff, G. and M. Johnson (1999), *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, New York: Basic Books.

Levshina, N. (2016). When variables align: A Bayesian multinomial mixed-effects model of English permissive constructions. *Cognitive Linguistics, 27*(2), 235–268

Lopez-Ozieblo, R. (2020). Proposing a revised functional classification of pragmatic gestures. *Lingua*, *247*. 1–18.

Marrese, O., Raymond, C., Fox, B., Ford, C., & Pielke, M. (2021). The Grammar of Obviousness: The Palm-Up Gesture in Argument Sequences. *Frontiers in Communication. 6*:663067.

McNeill, D. (1992). *Hand and mind. What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

McNeill, D. (2005). *Gesture and thought.* Chicago, IL: University of Chicago Press.

Merleau-Ponty, M. ([1962] 2005), *Phenomenology of Perception*, trans. C. Smith, New York: Taylor & Francis.

Mendoza-Collazos, J., & Zlatev, J. (2022). A cognitive-semiotic approach to agency: Assessing ideas from cognitive science and neuroscience. *Biosemiotics, 15*(1), 141-170.

Mittelberg, I. (2017). Embodied frames and scenes: body-based metonymy and pragmatic inferencing in gesture. *Gesture 16*, 204–233.

Mittelberg, I., & Joue, G. (2017). Source actions ground metaphor via metonymy: towards a frame-based account of gestural action in multimodal discourse, In B. Hampe (Ed.). *Metaphor: Embodied Cognition and Discourse*, (pp. 119–137). Cambridge: Cambridge University Press.

Möttönen, T. (2016), *Construal in Expression: Intersubjective Approach to Cognitive Grammar*, Helsinki: University of Helsinki.

Mouratidou, A. (in press). Investigating choice awareness through cognitive semiotics. In A.

Bilgari (Ed.) *Open Semiotics Vol 5*. Editions L'Harmattan.

Mouratidou, A., Zlatev, J., & van de Weijer, J. (2024). The body says it all: Non-verbal indicators of choice awareness. *Cognitive Semiotics, 17*(2), 233–266.

Moya, P. (2014). Habit and embodiment in Merleau-Ponty, *Frontiers in Human Neuroscience*, 8:542

Müller, C. (2004). Forms and uses of the Palm Up Open Hand. A case of a gesture family? In C. Müller & R. Posner (Eds.), *Semantics and pragmatics of everyday gestures* (pp. 234–256). Berlin, Germany: Weidler.

Müller, C. (2014). Gesture as "deliberate expressive movement". In M. Seyfeddinipur & M. Gullberg (Eds.), *From gesture in conversation to visible action as utterance,* (pp. 127–152). Amsterdam: John Benjamins.

Müller, C. (2017). How recurrent gestures mean: conventionalized contexts-of-use and embodied motivation. *Gesture*, *16*, 278–306

Müller, C., Bressem, J., & Ladewig, S. (2013). Towards a grammar of gesture: A form-based view. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & S. Tessendorf (Eds.), *Body - language - communication. An international handbook on multimodality in interaction (38.1)* (pp. 707–733). Berlin: De Gruyter Mouton.

Payrató, L., & Teßendorf, S. (2014). Pragmatic gestures. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill & J. Bressem (Eds.), *Body, language, communication: An international handbook on multimodality in human interaction (38.2)* (pp. 1531–1539). Berlin & Boston: De Gruyter Mouton.

Peirce, C. S. (1931–1958). *Collected Papers I-VIII*. [Hartshorne, C.; Weiss, P.; Burks, A. (eds.)] Cambridge: Harvard University Press [In-text references given as CP, followed by volume and paragraph numbers].

Pelkey, J. (2023). Embodiment and language. *WIREs Cognitive Science, 14*(5), e1649.

Pielli, L., & Zlatev, J. (2020). The cyborg body: Potentials and limits of a body with prosthetic limbs. *Cognitive Semiotics, 13*(2), 1–30.

Ruth-Hirrel, L. (2018). *Cyclic gestures and multimodal symbolic assemblies: an argument for symbolic complexity in gesture*. [Doctoral Thesis] University of New Mexico.

Sokolowski, R. (2000). Introduction to Phenomenology. Cambridge University Press.

Sonesson, G. (2007), From the meaning of embodiment to the embodiment of meaning: A study in phenomenological semiotics, In T. Ziemke, J. Zlatev and R. Frank (Eds), B*ody, Language and Mind. Vol 1: Embodiment*, (pp. 85–128), Berlin: Mouton de Gruyter.

Sonesson, G. (2010). From mimicry to mime by way of mimesis: Reflections on a general theory of iconicity, *Sign Systems Studies, 38*(1–4), 18–65.

Sonesson, G. (2012). The foundation of cognitive semiotics in the phenomenology of signs and meanings. *Intellectica, 58*(2), 207-239

Sonesson, G. (2021). The relevance of the encyclopaedia. From semiosis to sedimentation and back again. In N. Drăgan (Ed.), *Differences, Similarities and Meanings: Semiotic Investigations of Contemporary Communication Phenomena*, (pp. 97-120). Berlin, Boston: De Gruyter Mouton.

Stampoulidis, G., Bolognesi, M., & Zlatev, J. (2019). A cognitive semiotic exploration of metaphors in Greek street art. *Cognitive Semiotics, 12*(1), 1–20.

Uhrig, P. (2021). *Large-scale multimodal corpus linguistics: The big data turn*. [Habilitation Thesis]. Erlangen: Friedrich-Alexander-Universität Erlangen-Nürnberg.

Winter, B., & Bürkner, P. (2021). Poisson regression for linguists: A tutorial introduction to modelling count data with brms. *Language & Linguistics Compass*, e12439.

Woelert, P. (2011). Human cognition, space, and the sedimentation of meaning. *Phenomenology and the Cognitive Sciences 10*(1), 113–137.

Zahavi, D. (2003), *Husserl's Phenomenology*, Stanford: Stanford University Press.

Zahavi, D. (2018). *Phenomenology: The basics*. Routledge.

Zima, E., & Bergs, A. (2017). Multimodality and construction grammar. *Linguistics Vanguard, 3*(s1), 1–9.

Zlatev, J. (2005). What's in a schema? Bodily mimesis and the grounding of language. In B. Hampe (Ed.), *From perception to meaning: Image schemas in cognitive linguistics* (pp. 313–343). Berlin & New York: Mouton de Gruyter.

Zlatev, J. (2007). Language, embodiment and mimesis. In T. Ziemke, J. Zlatev & R. Frank (Eds.), *Body, language, mind. Vol 1: Embodiment*, (pp. 297–337). Berlin: de Gruyter.

Zlatev, J. (2008). The coevolution of intersubjectivity and bodily mimesis. In J. Zlatev (Ed.), *The shared mind: Perspectives on intersubjectivity*, (pp. 215–244). Amsterdam: Benjamins.

Zlatev, J. (2009). The Semiotic Hierarchy: Life, Consciousness, Signs and Language. *Cognitive Semiotics, 4*, 169–200.

Zlatev, J. (2010), Phenomenology and Cognitive Linguistics, In D. Schmicking and S. Gallagher (Eds.), *Handbook of Phenomenology and Cognitive Science*, (pp. 415–443), Berlin: Springer.

Zlatev, J. (2012). Cognitive semiotics: An emerging field for the transdisciplinary study of meaning. *The Public Journal of Semiotics*, *4*(1), 2–24.

Zlatev, J (2014). Image schemas, mimetic schemas and children's gestures. *Cognitive Semiotics, 7* (1), 3–29.

Zlatev, J. (2015a). Cognitive semiotics. In P. Trifonas (Ed.), *International handbook of semiotics*, (pp. 1043-1067). New York: Springer.

Zlatev, J. (2015b). The emergence of gestures. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence*, (pp. 458–477). New York: Wiley.

Zlatev, J. (2016). Turning Back to Experience in Cognitive Linguistics via Phenomenology, *Cognitive Linguistics, 27*(4), 559–572.

Zlatev, J. (2017). Embodied Intersubjectivity. In B. Dancygier (Ed.), *The Cambridge Handbook of Cognitive Linguistics* (pp. 172–187). Cambridge: Cambridge University Press.

Zlatev, J. (2018). Meaning Making from Life to Language: The Semiotic Hierarchy and Phenomenology. *Cognitive Semiotics, 11*(1), 1–18.

Zlatev, J. (2023). The Intertwining of Bodily Experience and Language: The Continued Relevance of Merleau-Ponty, *Histoire Épistémologie Langage*, *45*(1), 41–63

Zlatev, J. (2024). Constraining metaphor and metonymy in language and depiction: A cognitive semiotics approach. *Studies in Logic, Grammar and Rhetoric, 69*(1), 7-29.

Zlatev, J. (in press). Five pillars of cognitive semiotics. In A. Bilgari (Ed.) *Open Semiotics Vol 5.* Editions L'Harmattan.

Zlatev, J. & Andrén, M. (2009). Stages and transitions in children's semiotic development. In J. Zlatev, M. Andrén, C. Lundmark, & M. J. Falck (Eds.), *Studies in Language and Cognition*, (pp. 380–401). Newcastle: Cambridge Scholars

Zlatev, J., & Blomberg, J. (2016). Embodied intersubjectivity, sedimentation and non-actual motion expressions. *Nordic Journal of Linguistics, 39*(2), 185-208.

Zlatev, J., & Blomberg, J. (2019). Norms of language: What kinds and where from? Insights from phenomenology. In A. Mäkilähde, V. Leppänen, & E. Itkonen (Eds.), *Normativity in language and linguistics*, (pp. 69–101). Amsterdam: John Benjamins.

Zlatev, J., & Konderak, P. (2022). Consciousness and Semiosis. In J. Pelkey (Ed.), *Bloomsbury Semiotics Volume 1: History and Semiosis*, (pp. 169–191). Bloomsbury Academic.

Zlatev, J., & Möttönen, T. (2023). Cognitive linguistics and semiotics. In J. Pelkey & P. Cobley (Eds.), *Bloomsbury semiotics: Volume 4: Semiotic movements* (pp. 269–292). Bloomsbury Academic.

Zlatev, J. & Mouratidou, A. (2024). Extending the life world: Phenomenological triangulation along two planes. *Biosemiotics. 17*(2). 407–429.

Zlatev, J., Żywiczyński, P., & Wacewicz, S. (2020). Pantomime as the original human-specific communicative system. *Journal of Language Evolution, 5*(2), 156–174.

Appendices

Appendix A – Questions used (English version)

		2HAS	other			2HAS	other	
	session 1	1 2 3	4 6 5		session 6	3 1 2	4 5 6	
	session 2	2 1 3	5 6 4		session 7	1 3 2	5 4 6	
	session 3	1 3 2	4 6 5		session 8	2 3 1	4 5 6	
	session 4	2 3 1	5 4 6		session 9	3 1 2	6 5 4	
	session 5	1 3 2	4 6 5		session 10	2 1 3	5 4 6	

Topic 1: Communication

Q1-1 What impact does the choice of medium (e.g., messaging app, phone, face-to-face) have on the quality of communication?

- The medium affects clarity and emotional tone; face-to-face allows for non-verbal cues, while messaging apps are faster but risk misinterpretation.

Q1-2 Do you tend to be more of a listener or a speaker during communication? Additionally, From the perspective that communication is the exchange of ideas and opinions, what is required for both the speaker and the listener?

- The speaker needs clarity, empathy, and engagement to convey ideas effectively, while the listener requires active listening, open-mindedness, and constructive feedback to ensure mutual understanding.

Q1-3 How can we adapt our communication style to better connect with individuals who have different personality type than you (e.g. introvert vs. extrovert)?

- By observing their preferences; for introverts, use thoughtful and concise language, while with extroverts, engage in active dialogue and energy-matching.

Topic 2: Travel and transportations

Q2-1 Do you prefer to stay going back and forth between familiar places, like school and home, or do you enjoy moving to various locations in search of new environments?

- A mix of both is ideal—I like familiarity for stability but also enjoy occasional adventures to new places.

Q2-2 How does the experience of travel differ depending on the mode of transportation (e.g., train, plane, car, bicycle), especially when you travel between two or more locations?

- Traveling by train is relaxing, while planes are faster but a bit more stressful; cars offer flexibility, and bicycles give a sense of freedom and physical activity.

Q2-3 What are the future implications of autonomous vehicles on back-and-forth commuting and overall transportation systems?

- Autonomous vehicles may transform transportation systems by improving traffic flow, reducing fuel consumption, and cutting commute times, but they also require major infrastructure upgrades.

Topic 3: Different kinds of relationships

Q3-1 What are important in keeping any kinds of relationships (family, friend, partner, colleague, etc.) stable? Do they differ depending on the kinds of relationships?

- Trust, respect, and communication are key in all relationships, but their expression differs; for example, colleagues require adequate distance, while partners/families need emotional intimacy.

Q3-2 How do friendships made in childhood differ from those formed in adulthood, and what are the reasons behind these differences?

- In childhood, friendships are carefree and focused on fun, while adult friendships are often deeper, built on trust and understanding, shaped by life experiences.

Q3-3 What are the unique challenges of hierarchical relationships, such as teacher-student or manager-employee relationships?

- Hierarchical relationships can discourage open dialogue, leading to miscommunication or unmet expectations.

Topic 4: Changes

Q4-1 What are some strategies to cope with major life changes, such as moving, career shifts, or relationship transitions?

- Practicing self-care, seeking professional advice, and setting realistic goals can reduce stress during major shifts.

Q4-2 How have social norms and values shifted over the past century, and what might be driving these changes?

- Social norms have shifted toward inclusivity and equality, driven by technological advancements, globalization, and awareness of human rights.

Q4-3 How does the process of aging change our priorities, goals, and perspectives on life?

- With time, perspectives broaden; challenges are seen as life lessons, and personal contentment becomes more important than fulfilling societal expectations.

Topic 5: Habits and routines

Q5-1 Do you have any habits or routines that you've repeated over and over again throughout your life?

- Reading for at least 30 minutes before bed has been a routine that helps me relax and learn continuously.

Q5-2 How can habits and routines help in achieving long-term goals? Have you had such experiences?

- Habits like writing a daily to-do list helped me stay on track for a major project.

Q5-3 How do routines and habits differ between workdays and weekends, and how can one maintain balance?

- Workday routines are structured and productivity-focused, while weekend routines are more relaxed and centered around leisure or hobbies.

Topic 6: Happiness

Q6-1 What does happiness mean to you, and how has your definition of happiness

evolved over time?

- Happiness used to mean constant excitement and fun, but now it's about stability, meaningful connections, and personal growth.

Q6-2 What are the differences between short-term pleasure and long-term happiness, and how can we balance both?

- Short-term pleasure, like eating favorite food, provides instant satisfaction, while long-term happiness stems from fulfilling goals and nurturing relationships. Balancing both involves enjoying small joys without losing bigger aspirations.

Q6-3 How does comparison to others impact happiness, and what are strategies to minimize negative comparisons?

- Comparison often diminishes happiness by creating unnecessary pressure; focusing on personal growth helps reduce its impact.

Note: I used ChatGPT as a source of inspiration for formulating the 18 specific questions and their example answers. More specifically, I asked ChatGPT to generate 20 questions for each topic. From these, I formulated three questions per topic, aiming to balance the content to avoid excessive redundancy while maintaining consistency. For the example answers, I asked ChatGPT to provide three responses to each question and then synthesized a single example answer based on those responses.

Appendix B – Questions used (Japanese version)

Topic 1: コミュニケーション

Q1-1 媒体の選択（例：メッセージングアプリ、電話、対面）はコミュニケーションの質にどのような影響を与えるでしょうか？

- メディアは明確さと感情的なトーンに影響を与えます。対面では非言語的なサインが伝わる一方、メッセージアプリは迅速ですが誤解のリスクがあります

Q1-2 コミュニケーションの際、あなたは聞き手と話し手どちらでいることが多いですか？また、コミュニケーションはアイデアや意見の交換という観点から、話し手と聞き手それぞれに必要なことはなんですか？

- 話し手には、アイデアを効果的に伝えるための明確さ、共感、そして関与が必要であり、一方で聞き手には、相互理解を確実にするための積極的な傾聴、柔軟な思考、建設的なフィードバックが求められます。

Q1-3 あなたと異なる性格タイプ（例: 内向的 vs 外向的）を持つ人とより良い繋がりを持つためには、どのようにコミュニケーションスタイルを適応させるべきですか？

- 相手の好みを観察することが大切です。内向的な人には慎重で簡潔な言葉を使い、外向的な人とは積極的な対話やエネルギーを合わせたコミュニケーションを心がけます。

Topic 2: 旅行・移動・交通

Q2-1 学校や家のような馴染みのある場所を行ったり来たりするのが好きですか、それとも新しい環境を求めて色々な場所に行くのが好きですか？

- 両方のバランスが理想的です。安定のためには馴染みのある場所が好きですが、新しい場所への冒険も時々楽しみます。

Q2-2 移動手段（例: 電車、飛行機、車、自転車）によって、複数の場所を移動する際の経験はどのように異なりますか？

- 電車はリラックスでき、飛行機は速いですが少しストレスがあり、車は柔軟性があり、自転車は自由さと身体的な活動を感じられます。

Q2-3 自動運転が通勤や通学などの行き来・そして交通システム全体に与える将来的な影響は何ですか？

- 自動運転車は、交通の流れの改善、燃料消費の削減、通勤時間の短縮によって交通システムを変革する可能性がありますが、大規模なインフラの整備も必要とされます。

Topic 3: 様々な人間関係

Q3-1 どのような人間関係（家族、友人、パートナー、同僚など）でも安定を保つために重要なことは何ですか？また、それは人間関係の種類によって異なりますか？

- 信頼、尊重、そしてコミュニケーションが全ての関係において重要ですが、その表現方法は異なります。例えば、同僚とは適度な距離が必要ですが、パートナーや家族とは感情的な親密さが必要です。

Q3-2 子供の頃に築かれる友達との関係は、大人になってから形成される友人関係とどのように異なりますか？また、その違いの理由は何ですか？

- 子供時代の友情は気楽で楽しいことに集中する一方で、大人の友情はしばしばより深く、信頼と理解に基づき、人生経験によって形作られることが多いです。

Q3-3 教師と生徒、または上司と部下のようなヒエラルキーのある関係における特有の課題は何ですか？

- ヒエラルキーのある関係は率直な対話を妨げる可能性があり、それが誤解や期待が満たされない状況につながることがあります。

Topic 4: 変化

Q4-1 引っ越し、キャリアの変化、関係の転換など、人生の大きな変化に対処するための方法は何ですか？

- 自己管理を実践し、専門的な助言を求め、現実的な目標を設定することで、ストレスを軽減できます

Q4-2 過去 1 世紀で社会的規範や価値観はどのように変化してきましたか？また、それらの変化を促している要因は何ですか？

- 社会的規範は包括性や平等に向けて変化してきました。この変化の要因には、技術革新、グローバル化、人権意識の高まりが挙げられます

Q4-3 歳を取るという過程で、優先事項、目標、人生観はどのように変化しますか？

- 時間が経つにつれ視野が広がり、挑戦が人生の教訓と見なされるようになります。また、社会的期待を満たすことよりも個人の満足が重要になります

Topic 5: 習慣・ルーティン

Q5-1 これまでの人生で何回も何回も繰り返してきた習慣やルーティンはありますか？

- 就寝前に少なくとも 30 分間読書をすることが習慣になっています。これはリラックスし、継続的に学ぶ助けになっています

Q5-2 習慣やルーティンは長期的な目標達成にどのように役立ちますか？そのような経験がありますか？

- 毎日やることリストを作成する習慣が、大きなプロジェクトを進める際に役立ちました。

Q5-3 平日と週末でルーティンや習慣はどのように異なりますか？また、どのように平日と週末のルーティンのバランスを保つことができますか？

- 仕事の日のルーティンは構造化され、生産性に焦点を当てています。一方で、週末のルーティンはリラックスしたもので、趣味や余暇に集中しています。

Topic 6: 幸せ・幸福

Q6-1 幸福とはあなたにとって何を意味しますか？また、その定義はどのように進化しましたか？

- かつて幸福は常に刺激的で楽しいものでしたが、今では安定、意味のあるつながり、個人的な成長が幸福の定義です。

Q6-2 短期的な喜びと長期的な幸福の違いは何ですか？また、それらのバランスをどのようにして取ることができますか？

- 短期的な喜び（例: 好きな食べ物を食べる）は即時的な満足をもたらしますが、長期的な幸福は目標を達成し、関係を育むことから生まれます。バランスを取るには、

小さな喜びを楽しみつつ、大きな目標を見失わないことが大切です。

Q6-3　他人と自分を比べることは幸せにどのような影響を与えますか？それによるネガティブな影響を最小限に抑えるための方法は何ですか？

- 比較は不要なプレッシャーを生み出し、幸福を減少させることがあります。個人的な成長に焦点を当てることで、その影響を軽減できます

Appendix C – Questions used (Swedish version)

Ämne 1: Kommunikation

Q1-1: Vilken påverkan har valet av medium (t.ex., sms, telefonsamtal, prata i verkligheten) på kvalitén av en konversation?

- Valet av medium påverkar tydlighet och emotionel ton; det är lättare med icke-verbal kommunikation i verkligheten (t.ex., minspel), medan sms är snabbare men kan lättare bli missuppfattade.

Q1-2: Föredrar du att lyssna eller att prata när det kommer till kommunikation? Dessutom, från perspektivet att kommunikation är ett utbyte av idéer och åsikter, vad krävs från både talaren och lyssnaren?

- Talaren behöver tydlighet, empati, och engagemang för att effektivt förmedla idéer, medan lyssnaren kräver aktivt lyssnande, öppensinnighet, och konstruktiv feedback för att säkerställa att man förstår varandra.

Q1-3: Hur kan man anpassa hur man kommunicerar för att få en bättre koppling med folk som har en annan personlighetstyp än en själv?

- Genom att observera deras preferenser; för introverta, använda eftertänksamt och koncist språk, medan med extroverta, engagera dig aktivt i dialogen och matcha deras energi.

Ämne 2: Resa och transport

Q2-1: Föredrar du att åka fram och tillbaka mellan bekanta platser, såsom ditt hem och din skola, eller föredrar du att åka till nya platser?

- En blandning av de två är perfekt – jag gillar bekanta platser för det känns stabilt, medan nya äventyr till nya platser också kan vara kul.

Q2-2: Hur varierar upplevelsen av att resa beroende på om du reser via tåg, flygplan, bil, cykel, eller något annat?

- Att resa med tåg är avslappnade, flygplan är snabba men stressiga, medan bilar är flexibla, och cykel ger en känsla av frihet och fysisk aktivitet.

Q2-3: Vad för påverkan kan självstyrande fordon ha på pendling fram och tillbaka och transportsystemet i dess helhet?

- Självstyrande fordon kan förändra transportsystemet genom att förbättra trafikflödet, att minska bränslekonsumtionen, och minska tiden det tar att pendla, men de skulle också kräva stora uppgraderingar i infrastrukturen.

Ämne 3: Olika typer av relationer

Q3-1: Vad är viktigt för att behålla relationer (vare sig när det kommer till familj, vänner, partner, kollegor, m.m.)? Varierar vad som är viktigt beroende på vilken typ av relation det är?

- Tillit, respekt, och kommunikation är viktiga i alla relationer, men hur man uttrycker det varierar; t.ex., gentemot kollegor behöver en viss distans, medan gentemot en partner och ens familj förväntas närhet.

Q3-2: Hur kan vänskaper som bildats när man är vuxen skilja sig från de som bildats när man var barn, och vad är några orsaker till dessa skillnader?

- När man är barn är vänskaper problemfria och fokuserade på att ha kul, medan vänskaper när man är vuxen brukar vara djupare och byggda på tillit och förståelse, samt är formade av livserfarenheter.

Q3-3: Vad är unika utmaningar med hierarkiska relationer, såsom mellan en student och en lärare, eller en anställd och en chef?

- Hierarkiska relationer kan hindra öppen kommunikation, vilket leder till missförstånd.

Ämne 4: Förändringar

Q4-1: Vad är några strategier för att hantera stora livsförändringar, såsom att flytta, byta karriär, eller stora förändringar i ens personliga relationer?

- Att ta hand om sig själv, att söka professionella råd, och att sätta realistiska mål kan hjälpa en att minska stress i stora förändringar.

Q4-2: Hur har sociala normer och värderingar förändrats under det senaste århundradet, och vad kan ha drivit dessa förändringar?

- Sociala normer har rört sig mer mot jämlikhet, vilket kanske har drivits av teknologiska framsteg, globalisering, och medvetenhet om mänskliga rättigheter.

Q4-3: Hur kan våra prioriteringar, mål, och perspektiv förändras när vi blir äldre?

- Ens perspektiv blir större när man blir äldre; utmaningar ses snarare som läxor, och att helt enkelt vara nöjd kan bli viktigare än att uppfylla samhällets förväntningar.

Ämne 5: Vanor och rutiner

Q5-1: Har du några vanor eller rutiner som du upprepat om och om igen under ditt liv?

- Att läsa i 30 minuter innan jag går och lägger mig har hjälpt mig slappna av, och har även förbättrat min inlärning.

Q5-2: Hur kan vanor och rutiner hjälpa oss att nå långsiktiga mål? Har du haft en sådan upplevelse?

- Vanor som att skriva att-göra listor varje dag har hjälpt mig hålla mig fokuserad för ett stort projekt.

Q5-3: Hur kan vanor och rutiner variera mellan vardag och helg, och hur kan man hålla en balans mellan de två?

- Rutinerna i vardagen är strukturerade och målet är fokus, medan rutinerna på helgen är mer avslappnade och målet är att slappna av.

Ämne 6: Glädje

Q6-1: Vad betyder glädje för dig, och hur har din definition av glädje utvecklats över tid?

- Glädje brukade betyda att vara upprymd och ha kul, men nu handlar det mer om stabilitet, betydelsefulla relationer, och personlig utveckling.

Q6-2: Vad finns det för skillnader mellan kortsiktiga nöjen och långsiktig glädje, och hur kan vi balansera dem?

- Kortsiktiga nöjen, som att äta ens favoritmat, ger omedelbar tillfredställelse, medan långsiktig glädje handlar mer om att uppnå mål och att ta hand om

relationer. Att balansera båda innebär att man njuter av små nöjen utan att förlora ens större drömmar.

Q6-3: Hur kan det påverka ens glädje att jämföra sig själv med andra, och vilka strategier kan man använda sig av för att minimera negativa jämförelser?

- Att jämföra sig med andra minskar ofta ens glädje eftersom det skapar onödig press; att fokusera på personlig tillväxt hjälper att minska dess påverkan.

Appendix D – Consent form

LUND UNIVERSITY

Informed consent

I have been made aware of the following information prior to participating in the study.

1.	**Background and purpose**	This study is part of the examination for the course SPVR01 (Master's thesis), at the Centre for Language and Literature, Lund University, with teacher/supervisor Prof. Jordan Zlatev. The purpose is to investigate the similarities and differences in multimodal communication (the combination of language, facial expressions, gaze, manual movements, posture, etc.) between Japanese and Swedish people.
2.	**The study**	The session starts with a short survey on language background. Followingly, the participant will be working in a pair, asked to talk about questions regarding various topics which are provided and facilitated by the researcher. It will take approximately 30-60 minutes to complete the session.
3.	**Handling and storing the data**	The session will be videorecorded by using an iPad/iPhone which is not connected to the internet. The recorded material will be transferred to an external hard drive and deleted from the iPad/iPhone immediately after the session is completed. All data (researcher's notes, answers to the questions, etc.) will be completely anonymous and there will be no way to trace these back to the participant. If the image of the participant is included in publications, the participant's face will be mosaiced to ensure anonymity. After the completion of this study, the data may be used for further research, with anonymity always strictly maintained.
4.	**Voluntary participation**	Participation is voluntary, and the participant can stop participating in the study at any time, even after the completion of the data collection session.
5.	**Responsible person**	**Daiya Kuryu (Author)**: daiya.kuryu@gmail.com Harry Polfeldt (supporting data collection from Swedish participants) Jordan Zlatev (Supervisor)
6.	**Compensation**	Upon completion of the study, the participant will receive a cinema ticket. The resulting MA thesis will be accessible at https://www.sol.lu.se/en/education/masters-programmes/languages-and-linguistics/student-papers/ around July 2025.

I confirm by my signature that I consent to participate. This form is made in two copies, one for me, and one for the responsible person.

Date and place

_____

Signature and clarification of signature

_____