Visual Communication. Theoretical, Empirical, and Applied Perspectives (ViCom) – Zentralprojekt

Prof. Dr. Cornelia Ebert, Johann Wolfgang von Goethe-Universität, Frankfurt am Main (coordinator)

Prof. Dr. Markus Steinbach, Georg-August-Universität, Göttingen

---

1 Aims and objectives of the Priority Programme

1.1 Anticipated total duration of the project

September 1, 2022 – August 31, 2028

First phase: September 1, 2022 – August 31, 2025

1.2 Aims and scope

The Priority Programme *Visual Communication. Theoretical, Empirical, and Applied Perspectives* (*ViCom*) investigates the specific features and linguistic significance of visual communication. This comprises sign languages as fully developed natural languages, which exclusively rely on the visual channel for communication, and visual means that enhance spoken language such as gestures. It aims at disclosing the characteristics of the visual modality as a communication channel and its interaction with other channels (especially the auditory channel) to develop a comprehensive theoretical linguistic model of human communication and its cognitive foundations. Although *ViCom* focuses on the investigation of sign languages and gestures as prime examples of visual human communication, the research agenda comprises the investigation of gestures in didactic and therapeutic contexts, their significance for language acquisition, gestural aspects of vocal communication and in the written modality, pictorial narratives as well as visual communication strategies of non-linguistic species and in multimodal human-computer interaction systems.

In recent theoretical linguistic research, both sign language and gestures have been analysed by applying the linguistic tools established for spoken languages. Only little attention has been paid to modelling the specific properties of visual transmission channels. However, it is now becoming evident that the formal linguistic repertoire needs to be extended to meet the specific requirements of visual communicative means. It is thus one of the main goals of *ViCom* to enhance the formal linguistic apparatus and to develop formally precise theories that can deal with and model visual and multimodal communication.

Although gesture research is a new and emerging topic in theoretical linguistics, it has been a central topic for decades in other disciplines, such as cognitive linguistics, semiotics, psychology, neurosciences, animal communication, cultural studies, and conversation analysis. Likewise, cognitive-oriented analyses of sign languages have long highlighted their modality-specific (visual) properties. From an applied perspective, the need for natural multimodal human-computer interfaces has been recognised for some time already. And finally, gestures are important for didactic and therapeutic uses due to their well-established role in language acquisition and learning. So, while there already exists a substantial body of empirical data and a considerable amount of insights into the form and functions of visual communication, this knowledge is distributed across disciplines and has not yet been collected in a unified and comprehensive way. *ViCom* as a highly interdisciplinary Priority Programme aims at bringing together the above-mentioned disciplines and research traditions under a theoretical, empirical, and applied perspective and will use new insights generated in the projects to develop a new comprehensive theoretical model of language and communication that is able to account for all modalities. It will also help to advance linguistic methodology in technological, therapeutic, and didactic environments.

With formal linguistic semantic and pragmatic projects building the core of *ViCom*, the Priority Programme is designed to secure the interaction of projects with a formal linguistic focus and projects from those disciplines that can look back on a longstanding tradition of research in

gesture, sign, and other phenomena of visual communication. Hence, in addition to pragmaticists, semanticists, and philosophers of language, the expertise of cognitive linguists, semioticians, psycholinguists, neuroscientists and computer scientists, as well as learning and education researchers is indispensable for the overall success of this enterprise.

The *Zentralprojekt* provides the research infrastructure for individual projects within *ViCom*. The budget requested covers the expenses for coordination and components of the research infrastructure which are shared across all projects participating in the Priority Programme and which therefore need to be centrally administered. *ViCom* consists of about 15-18 individual research projects in each of the two funding periods (2022–25, 2025–28). Overall, we expect about 60-80 directly participating researchers at any time. In addition, we will establish working relations with other research projects in Germany where this is beneficial for the goals of *ViCom*. Researchers in such associated projects will also be able to participate in events and activities of *ViCom*. Furthermore, *ViCom* involves international partners including project PIs coming from non-German institutions that are international experts on visual communication and internationally renowned Mercator fellows. Another important aspect of the project aims at communicating research aims and findings beyond academic research communities and to the public at large. With a balanced catalogue of measures, *ViCom* also aims at ensuring the success of early career researchers and securing equal opportunities, especially for women and deaf researchers.

1.3 Objectives and programme

In *ViCom*, visual communication phenomena are investigated in **three basic dimensions**: (a) theoretical, (b) empirical, and (c) applied. Additionally, all projects are organised in at least one of **six project clusters** corresponding to the different research objects relevant for *ViCom*: (1) gesture, (2) sign language, (3) visual studies, (4) animal communication, (5) didactic and clinical aspects and language acquisition, and (6) human-computer interaction systems. In the following, we first describe the aims of the six project clusters before outlining the scope of the three basic dimensions.

1.3.1 Project clusters

ViCom aims at further developing linguistic theories by involving projects from each of the six clusters to advance our understanding of visual phenomena that are relevant for communication. As core visual components of interhuman communication, the first two clusters, gesture and sign, build the center of the programme.

(1) Gesture: In contrast to spoken and sign languages, speech-accompanying gestures (comprising manual gestures, facial expressions, body postures, brow movements, eye gazes, etc.) are not an autonomous linguistic system. They are communicative means that contribute significant information to a statement in addition to the linguistic content of the spoken material. Speech and gesture jointly contribute to the core meaning of an utterance (Kendon 1980; McNeill 1992; Liddell 2003; Kita & Özyürek 2003; Fricke 2012; Müller et al. 2013/2014). Traditionally, gesture studies have been conducted in semiotics, cognitive linguistics, cultural studies, and computational linguistics with a focus on the description and categorization of gestures, their role in language acquisition and their grammatical, semantic, and pragmatic interaction with spoken language in multimodal communication (McNeill 1992, 2005; Özçaliskan & Goldin-Meadow 2005; Kita 2009; Fricke 2012; Müller 2014; Lücking 2013; Harrison 2018; Ladewig 2020).

Gestures are typically more iconic than spoken language and thus semantically transparent to a certain degree. Furthermore, gestures are generally less conventionalised than words or signs. Recent formal linguistic models make different assumptions about how gesture and speech interact, linking the interaction to different linguistically well-known interaction patterns within speech (Kopp et al. 2004; Lascarides & Stone 2009; Giorgolo 2010; Lücking 2013; Ebert & Ebert 2014; Alahvedzhieva et al. 2017; Schlenker 2018a; Esipova 2019; Rieser & Lawler 2020; Ebert et al. 2020, cf. also the discussion in Hunter 2019). Initial formalizations of this gesture-speech alignment vary in their predictions on how exactly timing influences interpretation (e.g. Schlenker

2018a, 2020 vs. Esipova 2019). Recently, first attempts have been made to empirically and experimentally validate some of the models developed to account for the interaction of gesture and language (cf. e.g. Esipova 2017; Brentari et al. 2018; Fenlon et al. 2018, 2019; Tieu et al. 2018; Ebert et al. 2020, to appear).

As for pointing gestures, they have always played a major role within the domain of philosophy of language and theoretical linguistics in the context of demonstratives (Kripke 1977; Kaplan 1989; Nunberg 1993 among many others). But although pointing gestures have been treated as an integral part of establishing the speaker's intended meaning in these theories, the act of pointing itself has not been further investigated in the traditional literature. Recently, Ebert & Ebert (2014), Lücking et al. (2015), and Ahn & Davidson (2017) have shown that some traditional analyses of gesture-speech combinations can be rethought by taking a closer look at the contribution of the gesture itself. Such a novel view has the potential to shed new light on classic phenomena and enlighten even persistent philosophical debates. This also applies to iconicity-based linguistic phenomena such as role shift and pointing in sign languages, where pointing signs are used to establish reference and anaphoric relations in discourse (Steinbach & Onea 2016).

Another area where gesture research connects to more traditional linguistic research is information and discourse structure. Ebert et al. (2011) show that certain units of gesture structures, i.e. so-called gesture phrases (Kendon 1980), align with focus phrases, an information-structural unit that marks new, important, or contrastive information (cf. also Dimitrova et al. 2016 on beat gestures and focus). Research has also shown that different types of iconic gestures can make the information status of a narration visible in spoken and sign languages (McNeill 1992; see also Kita & Özyürek 2003; Perniss 2007) and reveal the respective viewpoint or perspective of said narration (e.g. Stec 2012; Parrill 2010; Maier & Steinbach to appear; Ebert & Hinterwimmer to appear). Accordingly, so-called *character viewpoint gestures* (usually enactments with the whole body) convey central information and encode affective content. *Observer viewpoint gestures* (usually executed as manual gestures), on the other hand, are used more often when peripheral information is transferred (Beattie & Shovelton 2002; McNeill 1992; Parrill 2010). Studies have also demonstrated that gestures are organised according to topic-comment structures of speech (Kendon 1995; Seyfeddinipur 2004). Moreover, gestures make visible semantic threads of longer discourses and discourse topics (Enfield 2004), they mark different phases of a discourse such as the climax of a narrative (Ladewig & Hotze, in press) and enhance narrative performance (Vilà-Giménez & Prieto 2020). Gestures have also been shown to mark and visualise speech acts and sentence types (Streeck 2009; Bressem & Müller 2014; Cruz et al. 2017; Domaneschi et al. 2017; Harrison 2018; Brentari et al. 2018) as well as discourse moves and to regulate turn-takes in dialogues (Streeck & Hartge 1992; Schmitt 2005; Bohle 2007; Mondada 2007; Holler et al. 2018; Ginzburg & Lücking, to appear).

Cross-modally, certain spoken and written language phenomena can also be referred to as *gestures*. These are non-standard expressions or amendments from standard expressions that heavily draw on iconicity and have some depictive (rather than descriptive) character. This concerns *ideophones* (Dingemanse 2012; Ćwiek 2021) as well as *prosodic modulations* such as lengthening or pitch changes (Okrent 2002; Perlman & Benitez 2010) or their equivalents in written language (Fuchs et al. 2019). *ViCom* is also designed to investigate these vocal and written language gestures to be able to compare iconic means across different modalities, i.e. acoustic, graphemic, and visual gestures.

(2) Sign languages: Each theory of language and communication has to account for the fact that natural languages come in at least two modalities. Like spoken languages, sign languages are fully developed natural human languages. Unlike spoken languages, sign languages do, however, not use the auditive-oral modality but exclusively the visual-gestural modality. Since Stokoe's (1960) seminal work, sign languages have been intensively investigated and research on their structure and use has proven beyond any doubt that sign languages are complete and independent natural languages with a complex grammatical system (cf. Sandler & Lillo-Martin 2006;

Brentari 2010; Pfau et al. 2012; Goldin-Meadow & Brentari 2017). Likewise, psycho- and neurolinguistic studies on the acquisition, production, and processing of sign languages have shown that both modalities share the same neural networks and similar cognitive processes (Chen Pichler 2012; Corina & Spotswood 2012; Dye 2012; Hohenberger & Leuninger 2012).

Given these similarities, it is, however, equally important to identify the impact of modality on the architecture of linguistic structures and language use and to develop linguistic theories that account for both modality-independent as well as modality-specific properties. There appear to be three key modality-specific properties of sign languages which set them apart from spoken language. Firstly, sign languages systematically exploit their ability to express different grammatical categories simultaneously (Meier 2002; Aronoff et al. 2005). The second difference is due to the spatial nature of sign languages: sign languages use the geometrical properties of the three-dimensional signing space to realise morphosyntactic, semantic and pragmatic categories in a systematic way (Perniss 2012; Steinbach & Onea 2016). The third difference concerns the interaction of gesture and sign languages: sign languages not only have a gestural origin (Senghas & Coppola 2011; Sandler et al. 2014), but the close interrelation between manual and nonmanual gestures and sign language is still visible in the linguistic structures of developed sign languages and a core feature of visual-gestural communication (Emmorey & Reilly 1995; Rathmann & Mathur 2002; Liddell 2003; Goldin-Meadow & Brentari 2017; Schlenker 2018b). This third difference between the two modalities, i.e. the interaction between gesture and sign language, is the most relevant for *ViCom*. In sign languages, gestures can be used to express parts of the propositional content as, for instance, in role shift (Steinbach 2021), gestures can enter the lexicon (Janzen 2012), and gestures can be grammaticalised and pragmatized (van Loon et al. 2014).

Although these examples show that sign languages draw on gestural sources and systematically integrate gestural components at different levels of grammar and language use, the impact of gestures on the grammar of sign languages and the extent to which gestures and sign language interact remains a controversial topic in sign language linguistics (see e.g. Liddell 2003; Wilbur 2013; Vigliocco et al. 2014; Goldin-Meadow & Brentari 2017; Brentari & Goldin-Meadow 2017). Formal linguistic analyses have only recently addressed the relevance of gestures and thus integrated insights from cognitive linguistics into formal analyses, which is a promising new development. However, a comprehensive model of the similarities and differences between gesture and sign language and the systematic interaction of gesture and sign language framed within more recent formal linguistic theories has yet to be devised (for first analyses see Davidson 2015; Schlenker 2018a,b; Ebert 2018; Maier 2018; Hübl et al. 2019; Oomen 2020; Maier & Steinbach 2022).

A related difference between spoken and sign languages, which is important for the discussion of the impact of modality, relates again to the visible nature of sign languages. For years, the prevailing opinion in the literature has been that sign languages are ‘more iconic’ than spoken languages. A potential explanation for this difference is often seen in their visual basis, i.e. in the fact that the visual system allows for more flexibility in iconic mappings than the auditive system (Perniss et al. 2010; Taub 2012). Whether this is actually the case is still under discussion and up for investigation (Perlman & Lupyan 2018).

(3) Visual studies: In addition to human face-to-face communication, there are other forms of culturally grounded visual communication that have recently attracted the attention of theoretical linguists. These are for instance non-verbal pictorial representations (Abusch 2013; Abusch & Rooth 2017; Greenberg 2018), picture narratives and comics (Maier 2019) as well as films (Cumming et al. 2017), certain forms of dance (Napoli & Kraus 2017; Patel-Grosz et al. 2018; Charnavel 2019), and the use of emojis in social media and elsewhere (Grosz et al. 2021; Kaiser & Grosz 2021). All these communication forms make heavy use of iconic symbols. A theory of visual communication that aims at representing and seriously integrating a component of iconic semantics will evidently profit from investigating these other forms of communication that are based on iconic mappings of meaning.

Additionally, first attempts have been made to analyse visual narratives, i.e. films, with their inherent camera viewpoint rules and natural interpretation schemes with the help of formal semantic methods (Cumming et al. 2017). Films are, like pictorial narratives, a sequence of discontinuous parts, in this case, camera shots (cf. Schlenker 2019). Like in pictorial narratives, we find preferences for the interpretation of viewpoints and viewpoint shifts in films, which are well documented, but not well understood with respect to the question of where they stem from.

(4) Animal communication: Recent developments have led some researchers in formal semantics to begin to investigate animal communication with established formal linguistic tools, sometimes in collaboration with anthropologists and neurobiologists. So far, the majority of studies on animal communication systems with a semantic focus have mainly explored auditory-based systems. The aim of these new models for animal communication is to analyse empirically gathered data with formal linguistic tools. Furthermore, a semantic analysis of animal communication strategies is put on a truth-conditional basis with the relevant question being: under what conditions is a certain communication strategy (e. g. an alarm call) 'true' or appropriate?

ViCom takes a more holistic view on animal communication systems by focusing on gesture-based as well as multi-modal communication. For instance, great apes, like humans, combine visual and non-visual means to communicate, i.e. alarm calls, manual and facial gestures, body postures, etc., which provides an opportunity for the interaction of visual and non-visual communication strategies and interactions (see Fröhlich et al. 2019; Hobaiter et al. 2017). Furthermore, Pika (2008) and Kersken et al. (2018) showed that there is a great overlap between the repertoire and types of gestures used by great apes and those used by pre-linguistic human children, although the difference between the gestures of humans and nonhuman apes is often stressed in the literature (e.g. Tomasello 2006; see also Liebal et al. 2005 and Krause et al. 2018).

A closer look at visual strategies of nonhuman animal communication also has the potential to shed new light on the prevalent question about the origin of human language. While we know that sign languages have gestural origins (Senghas & Coppola 2011; Sandler et al. 2014), this is only one of the theories for the origin of spoken languages (Tomasello 2008).

(5) Didactic and clinical aspects & language acquisition: The eminent role of gestures in (spoken and sign) language acquisition (Özsaliskan & Goldin-Meadow 2005; Morgenstern & Guidetti 2017; Rohlfing 2019), cognitive development (Church & Goldin-Meadow 1986; Goldin-Meadow & Alibali 2013), and for conscious learning is well-established. On basis of experimental studies, Goldin-Meadow and her colleagues have shown that gestures can reveal whether children are in a transition phase to reach a new cognitive-developmental step. Furthermore, gestures are an important precursor of children's subsequent language development (see Colonnesi et al. 2010; Goldin-Meadow & Alibali 2013; Beuker et al. 2013; Igualada et al. 2015; Özçalışkan et al. 2016). In addition, it has been shown that the early use of pointing gestures plays a role in predicting language development of clinical populations, such as children with language delay (Lüke et al. 2020) or autism spectrum disorder (Özçalışkan et al. 2016; Sansavini et al. 2019). It is also well-known from the literature that vocabulary learning can be enhanced significantly with manual gestures (e.g. Capone Singleton 2012; Kushch et al. 2018). Producing gestures, as well as watching someone else produce gestures while learning new words, helps learners memorize said words and iconic congruent gestures have been shown to have a more significant effect than meaningless or incongruent ones (see e.g. Kelly et al. 2009; Macedonia et al. 2011; Andrä et al. 2020). This beneficial effect of iconic gestures for word learning was also found in children with developmental language disorder (Weismer & Hesketh 1993; Lüke & Ritterfeld 2014; Vogt & Kauschke 2017; van Berkel-van Hoof et al. 2019) and is now under investigation as a means to support children with language barriers in inclusive settings (BMBF project, 01UL1811X).

During the last decade, gestures have also been shown to be important for learning in mathematics (Goldin-Meadow et al. 2009; for learning anatomical knowledge, see Cherdieu et al. 2017). Studies range from investigations on early childhood mathematics education (Elia 2018) to adult learning during the display of mathematical content, e.g. the concept of fractions (Edwards 2009).

Vogel & Huth (2020) stress the importance of actions and gestures for learning mathematical concepts and point to their relevance for digital learning environments, with touch gestures in digital settings triggering digital actions on the screen. To the same end, it has been shown that vocabulary learning in a digital learning setting involving a congruent finger gesture that controls an animated figure on a touch device could be enhanced by using this interactive learning method (Ebert et al. 2018). *ViCom* aims at further investigating the supporting role of gestures in language acquisition and educational settings in order to gain more insights into their nature and the significance of gestures for language.

(6) Multimodal human-computer interaction and computational methods: Visual communication means have a longstanding tradition in the area of human-computer interaction. To build artificial agents that are capable of natural, human-like interaction, these agents have to use verbal as well as non-verbal communication means. Most prominently, this led to the development of *Embodied Conversational Agents* (ECAs; Cassell 2000; Kopp et al. 2004; Pelachaud 2005). Continuing these technical developments, ECAs have been used as social agents, emphasising their emotional display and facial expressions (Kopp 2010; Krenn et al. 2011). An especially interesting and important area of application are automatic translation and avatar technologies which are not able to be used in linguistic and corpus-based research but can also provide accessible communication in multimodal contexts, especially for sign language communication (Wolfe et al. 2016). Within virtual reality (VR) applications, gestures are an input device that serves to steer the application (as such they are an alternative to controller-based VR applications). Gesture recognition can be sensor-based, glove-based, or vision-based (e.g., Li et al. 2019). It consists of mapping manual input onto predefined “gesture lexicons” like the by now ubiquitous touch gestures for smartphones and other new applications (e.g., Kamel Boulos et al. 2011).

1.3.2 Basic dimensions

While *ViCom*’s orientation of the content relates to the six clusters described in the previous section, methodologically the programme makes reference to three basic dimensions, (a) theoretical, (b) empirical, and (c) applied, that will be laid out in this section and that define the overarching objectives of this programme.

(a) Theoretical model: The first major goal of *ViCom* is the development of new cross-modal formal theories of linguistic structure and communication. To this end, existing theories e.g. from theoretical linguistics, sign language linguistics, or gesture theory, need to be compared, modified and advanced to capture visual communication phenomena. This requires the development of new formal tools and an extension of the formal apparatus to be able to describe, for example, iconic, demonstration-based and sociolinguistic components in the same formally rigorous way as other language phenomena. The final result should not only explain gestural meaning contributions and the systematic interaction of gestures with sign languages but also account for visually based cultural communication forms (as in films or comics or with the use of emojis in social media). Furthermore, it should provide a novel view on animal communication systems and language development. To achieve this aim, it is crucial that linguists collaborate closely with neuroscientists, anthropologists, and researchers from biocognition, language acquisition, animal studies, and other research areas laid out above.

As pointed out above, *ViCom* starts from a theoretical linguistic point of view, namely that of semantics and pragmatics or in other words the linguistic research tradition which develops linguistic theories on the basis of formal and mathematical models. Within this research tradition, the existing formal models of language meaning and use have, until very recently, been based solely on spoken or written language phenomena. Nonetheless, visual communication systems differ in several aspects from spoken language and the analysis of the visual modality requires modified or even new formal tools, which will differ from the existing ones. For example, sign languages

and gestures have iconic properties, i.e. they map aspects from the real world iconically to the visual space. In addition, gestures are used to demonstrate (aspects of) actions and events visually (Clark & Gerrig 1990). Iconicity and demonstration are two phenomena that are crucially different from other linguistic phenomena that are well-examined. New formal models and theories will be needed to do justice to these specific properties of visual communication. This will also shed new light on the analysis of spoken languages (cf. Schlenker 2014, 2019). The co-existence of arbitrariness and iconicity in language is now also an emerging topic of discussion in spoken and written language linguistics (Dingemanse et al. 2015; Lupyan & Winter 2018; Fuchs et al. 2019; Krivochen & Lacková 2020).

(b) Empirical coverage: A second major goal of *ViCom* is the collection and unification of empirical and experimental data from various disciplines, the advancement of available methods and the development of novel methodology for data collection and validation of formal models in the domain of visual communication. In recent years, researchers have started to build and analyse representative multimodal corpora of sign languages as well as gesture-speech interaction (Joo et al. 2017; Paggio et al. 2020). Two representative corpora important for *ViCom* are the *DGS Korpus*, a long-term corpus project on German Sign Language (DGS) of the Academy of Sciences in Hamburg (Hanke 2016), and the *Bielefeld Speech and Gesture Alignment Corpus (SaGA)* (Lücking et al. 2013). Both corpora are among the largest corpora in their field and have set new standards for the collection of multimodal data and the linguistic annotation of sign languages and speech accompanying gestures. Notable other corpora are the *Multimodal SmartKom Corpus* (Schiel et al. 2002), the *AMI meeting corpus* (Carletta et al. 2006), DUEL (Hough et al. 2016), FIGURE (Lücking et al. 2016), and *The Corpus of Interactional Data* (Blache et al. 2017). To annotate these multimodal corpora, video-based annotation tools are used (Cassidy & Schmidt 2017). *ViCom* aims to further develop such multimodal corpora, for example the integration of motion tracking technologies (motion capture, open pose) and defining common standards for multimodal annotations.

Likewise, new developments in experimental sign language and gesture research offer new perspectives on visual communication through the investigation of psycho- and neurolinguistic aspects of the acquisition, production, and comprehension of sign, speech, and gesture in different contexts, using various experimental methods such as controlled production and judgment tasks, motion capture, eye tracking, ERP, and fMRI. Recent experimental studies have investigated different topics and have focussed on the interaction of spoken language and gestures (de Ruiter 2000; Kita & Özyürek 2003; Mittelberg 2018; Tieu et al. 2018; Ebert et al. 2020, to appear), sign language and gestures from various perspectives (Emmorey et al. 2003; Schouwstra 2012; Hellbernd & Sammler 2016; Weisberg et al. 2016; Finkl et al. 2020; Leonard et al. 2020), and the impact that visual input has on linguistically relevant categories such as *alternatives* or *presupposition* (Tieu et al. 2019; Schlenker 2021).

An important aspect of empirical studies on multimodal communication is the advancement of methods, tools, and common standards for experimental studies, data annotation, and data evaluation (Van Herreweghe & Vermeerbergen 2012; Holler 2014; Wagner et al. 2014; Orfanidou et al. 2015; Wei et al. 2016; Pouw et al. 2020). Working with visual data and stimuli captured on video and eliciting new data beyond classical grammaticality and felicity judgments requires the development of new integrated methods in psycho- and neurolinguistics as well as in computer science to capture the specific properties of multimodal data in experimental and corpus studies. Likewise, new standards of image processing, (automatic) annotation, and statistical evaluation are necessary to handle multimodal corpus data. In this respect, *ViCom* can build on computational linguistic and experimental expertise gained in the SaGA and DGS corpus projects, the Priority Programme XPRAG.de, and the experience within various labs that will be involved in *ViCom*. Empirical research on sign languages additionally faces the challenge that deaf participants are a special linguistic and cultural minority with specific linguistic and social features in many societies (Quer & Steinbach 2019). Therefore, for ethical reasons, any kind of empirical

investigation has to consider the specific needs of deaf participants.

(c) Application: The third major goal of *ViCom* is to learn more about the possibility of transferring results from research on actual physical gesturing to the applied use of gesturing in digital learning settings or virtual world environments. One pertinent question is the following: can gestures in these environments enhance learning in the same way as we know that actual physical gestures support vocabulary learning or mathematical learning? *ViCom* also aims at finding out more about the key role that gesture, action, and visual means play for language development and learning in different environments, such as pre-schools or schools, and for different groups (people with communicative barriers, typically developing children, people in the autism spectrum, children with language delay, and others). It will furthermore foster the development of new didactic theories and tools and aims to contribute to improving the utilization of sign language and gestures in the different settings (e.g. within schools, interpreter training, vocational training, etc.). Additionally, *ViCom* aims at facilitating the development of automatic annotation programmes, gesture and sign recognition systems, as well as sign and gesture generation systems (multimodal avatars) and human-computer interaction systems. One important question is how human-computer interaction can be made more natural for users.

It is essential for the success of this programme that the envisaged clusters and dimensions involve different disciplines and foster exchanges between these fields and their different research traditions, backgrounds, and methods, as only such a diverse and interactive interdisciplinary constellation will bring the merits that we expect. *ViCom* will then pave the way for new long-term collaborations and the establishment of new research groups, thus effectively shaping research on gesture, sign language, and other forms of visual communication in Germany and internationally.

The structure of the Priority Programme ensures optimal coherence across projects and facilitates close collaborations between different projects in various respects. Projects within one cluster will work in close collaboration, delegate one PI to the steering committee, apply for joined workshops in the workshop programme and propose Mercator fellows. To ensure a good interaction across the six potential clusters, projects in the same cluster will define common aims and exchange theoretical and methodological expertise, as well as interim results, in order to investigate all research objects with respect to all three dimensions. Likewise, projects with a focus on one of the three dimensions will collaborate to enhance theoretical models, empirical methods as well as technological, didactic, and therapeutic means. Projects with a strong empirical bias will establish a forum to discuss methodological aspects of different kinds of empirical research and to share recent developments in methodology. Specifically, we envisage a specific measure to ensure the exchange among data-oriented, i.e. empirically and experimentally focused or corpus-based projects: We will strive to build a network among *ViCom* members conducting empirical research and to develop a joint open science database with collectively developed shared annotation and coding schemes and common standards for metadata for audio-visual data. The network will be supported by an IT technician and led by one trained and renowned specialist among the PIs of *ViCom*. Members will meet in workgroups and workshops on a regular basis (for details see measure *open science database* below). Similarly, projects with a strong theoretical bias will discuss problems in theory building and share advances in the development of new theoretical models, predominantly via joint project workshops, short-term collaborations, or at the annual meetings. Finally, projects with an applied bias will be in close contact to test and promote the applicability of new findings in different areas. This can be accomplished by applied studies in varying contexts, the development of practical open-access tools, knowledge transfer, and third mission activities (see also the section on public relations below). Ideally, each project will be part of an interactive network of projects working on the same research object on the one hand and a network of projects investigating different topics, but using similar theories or methods or developing similar tools for application on the other hand.

1.4 References

Abusch, D. 2013. Applying discourse semantics and pragmatics to co-reference in picture sequences. *Proceedings of Sinn und Bedeutung* 17, 19-25.

Abusch, D. & M. Rooth. 2017. The formal semantics of free perception in pictorial narratives. *Proceedings of the 21st Amsterdam Colloquium*, 85-95.

Ahn, D. & K. Davidson. 2017. Where pointing matters: English and Korean demonstratives. *Presentation at NELS 2017*, Iceland. October 27-29.

Alahvedzhieva, K., A. Lascarides & D. Flickinger. 2017. Aligning speech and co-speech gesture in a constraint-based grammar. *Journal of Language Modelling* 5(3), 421-464.

Andrä, C., B. Mathias, A. Schwager, M. Macedonia & K. von Kriegstein, K. 2020. Learning foreign language vocabulary with gestures and pictures enhances vocabulary memory for several months post-learning in eight-year-old school children*. Educational Psychology Review* 32(3), 815-850.

Aronoff, M., I. Meir & W. Sandler. 2005. The paradox of sign language morphology. *Language* 81(2), 301-344.

Beattie, G., & H. Shovelton. 2002. An experimental investigation of some properties of individual iconic gestures that mediate their communicative power. *British Journal of Psychology* 93(2), 179-192.

Beuker, K.T., N.N.J. Rommelse, R. Donders & J.K. Buitelaar. 2013. Development of early communication skills in the first two years of life. *Infant Behavior and Development* 36(1), 71-83.

Blache, P., R. Bertrand, G. Ferré, B. Pallaud, L. Prévot & S. Rauzy. 2017. The corpus of interactional data: A large multimodal annotated resource. In Ide, N. & J. Pustejovsky (eds.), *Handbook of Linguistic Annotation*. Heidelberg: Springer, 1323-1356.

Bohle, U. 2007. *Das Wort ergreifen – Das Wort übergeben. Explorative Studie zur Rolle redebegleitender Gesten in der Organisation des Sprecherwechsels*. Berlin: Weidler Verlag.

Brentari, D. 2010. *Sign Languages. Cambridge Language Surveys*. Cambridge: CUP.

Brentari, D., J. Falk, A. Giannakiou, A. Herrmann, E. Volk & M. Steinbach. 2018. Production and comprehension of prosodic markers in sign language imperatives. *Frontiers in Psychology: Language Sciences* 9, 770, 1-14.

Brentari, D. & S. Goldin-Meadow. 2017. Language emergence. *Annual Review of Linguistics* 3, 363-388.

Bressem, J. & C. Müller. 2014. A repertoire of German recurrent gestures with pragmatic functions. In Müller, C., S. Ladewig, A. Cienki, E. Fricke, J. Bressem & D. McNeill (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction.* Berlin/Boston: De Gruyter Mouton, 1575-1592.

Capone Singleton, N. 2012. Can semantic enrichment lead to naming in a word extension task? *American Journal of Speech-Language Pathology* 21, 279-292.

Carletta, J., S. Ashby, S. Bourban, M. Flynn, M. Guillemot, T. Hain, J. Kadlec, V. Karaiskos, W. Kraaij, M. Kronenthal, G. Lathoud, M. Lincoln, A. Lisowska, I. McCowan, W. Post, D. Reidsma & P. Wellner. 2006. The AMI meeting corpus: A pre-announcement. In Renals S. & S. Benigo (eds.), *Machine Learning for Multimodal Interaction*. Heidelberg: Springer, 28-39.

Cassell, J. 2000. Embodied conversational interface agents. *Communications of the ACM* 43(4), 70-78.

Cassidy, S. & T. Schmidt. 2017. Tools for multimodal annotation. In Ide, N. & J. Pustejovsky (eds.), *Handbook of Linguistic Annotation*. Heidelberg: Springer, 209-227.

Charnavel, I. 2019. Steps towards a universal grammar of dance: Local grouping structure in basic human movement perception *Frontiers in Psychology* 10, 1364, 1-19.

Chen Pichler, D. 2012. Acquisition. In Pfau, R., M. Steinbach & B. Woll (eds.), *Sign Language. An International Handbook*. Berlin/Boston: De Gruyter Mouton, 647-686.

Cherdieu, M., O. Palombi, S. Gerber, J. Troccaz & A. Rochet-Capellan. 2017. Make gestures to learn: Reproducing gestures improves the learning of anatomical knowledge more than just seeing gestures. *Frontiers in Psychology* 8*,* 1689.

Church, R.B. & S. Goldin-Meadow. 1986. The mismatch between gesture and speech as an index of transitional knowledge. *Cognition* 23(1), 43-71.

Ćwiek, A. 2021. *Iconicity in language and speech*. PhD thesis, submitted at Humboldt University, Berlin.

Clark, H.H. & R.J. Gerrig. 1990. Quotations as demonstrations. *Language* 66, 764-805.

Colonnesi, C., G.J.J.M. Stams, I. Koster & M.J. Noom. 2010. The relation between pointing and language development: A meta-analysis. *Developmental Review* 30(4), 352-366.

Corina, D. & N. Spotswood 2012. Neurolinguistics. In Pfau, R., M. Steinbach & B. Woll (eds.), *Sign Language. An International Handbook*. Berlin/Boston: De Gruyter Mouton, 739-762.

Cruz, M., M. Swerts & S. Frota. 2017. The role of intonation and visual cues in the perception of sentence types: Evidence from European Portuguese variants. *Laboratory Phonology* 8(1), 1-23.

Cumming, S., G. Greenberg & R. Kelly. 2017. Conventions of viewpoint coherence in film. *Philosophers' Imprint* 17(1), 1-29.

Davidson, K. 2015. Quotation, demonstration, and iconicity. *Linguistics and Philosophy* 38(6), 477-520.

Dimitrova, D., C. Mingyuan, L. Wang, A. Özyürek & P. Hagoort. 2016. Beat that word: How listeners integrate beat gestures and focus in multimodal speech discourse. *Journal of Cognitive Neuroscience* 28(9), 1255-1269.

Dingemanse, M. 2012. Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass* 6(10), 654-672.

Dingemanse, M., D.E. Blasi, G. Lupyan, M.H. Christiansen & P. Monaghan. 2015. Arbitrariness, iconicity and

systematicity in language. *Trends in Cognitive Sciences* 19(10), 603-615.
Domaneschi, F., M. Passarelli, & C. Chiorri. 2017. Facial expressions and speech acts: Experimental evidences on the role of the upper face as an illocutionary force indicating device in language comprehension. *Cognitive Processing*, 1-22.
Dye, M.W.G. 2012. Processing. In Pfau, R., M. Steinbach & B. Woll (eds.), *Sign Language. An International Handbook*. Berlin/Boston: De Gruyter Mouton, 687-711.
Ebert, Ch., C. **Ebert** & R. Hörnig. 2020. Demonstratives as dimension shifters. *Proceedings of Sinn und Bedeutung* 24. Osnabrück, University of Osnabrück, 161-178.
Ebert, C. 2018. A comparison of sign language with speech plus gesture, commentary to Philippe Schlenker's target article 'Visible meaning: Sign language and the foundations of semantics'. *Theoretical Linguistics* 44(3-4), 239-249.
Ebert, C. & Ch. Ebert. 2014. Gestures, demonstratives, and the attributive/referential distinction, *Talk at Semantics and Philosophy in Europe* 7, ZAS, Berlin, June 2014. https://user.uni-frankfurt.de/~coebert/talks/CECE-SPE2014.pdf
Ebert, C., Ch. Ebert & R. Hörnig. 2018. Vocabulary learning improves with interactive finger gestures. In Rothkopf, C., D. Balfanz, R. Galuske, F. Jäkel, K. Kersting, J. Macke & B. Mohler (eds.), *14th Biannual Conference of the German Society for Cognitive Science* 11.
Ebert, C., S. Evert & K. Wilmes. 2011. Focus marking via gestures. *Proceedings of Sinn und Bedeutung* 15, 193-208.
Ebert, C. & S. Hinterwimmer. To appear. Free indirect discourse meets character viewpoint gestures: A reconstruction of Davidson's demonstration account with gesture semantics. To appear in *Linguistic Evidence 2020 Proceedings*.
Ebert, C., G. Pirillo & S. Walter. To appear. The role of gesture-speech alignment for gesture interpretation. To appear in *Linguistic Evidence 2020 Proceedings*.
Edwards, L.D. 2009. Gestures and conceptual integration in mathematical talk. *Educational Studies in Mathematics* 70, 127-141.
Elia, I. 2018. Observing the use of gestures in young children's geometric thinking. In Elia, I., J. Mulligan, A. Anderson, A. Baccaglini-Frank & C. Benz (eds.), *Contemporary Research and Perspectives on Early Childhood Mathematics Education*. Heidelberg: Springer, 159-182.
Emmorey, K., S. McCullough & D. Brentari. 2003. Categorial perception in American Sign Language. *Language and Cognitive Processes* 18(1), 21-45.
Emmorey K. & J. Reilly (eds.). 1995. *Language, gesture and space*. Hillsdale, NJ: Lawrence Erlbaum.
Enfield, N.J. 2004. On linear segmentation and combinatorics in co-speech gesture: A symmetry-dominance construction in Lao fish trap descriptions. *Semiotica* 149(1/4), 57-123.
Esipova, M. 2017. Contrast and distributivity in the semantics of alternation. *Proceedings of NELS 47*, 309-322.
Esipova, M. 2019. Composition and projection in speech and gesture. Unpublished PhD thesis. New York University. https://ling.auf.net/lingbuzz/004676
Fenlon, J., K. Cooperrider, J. Keane, D. Brentari & S. Goldin-Meadow. 2019. Comparing sign language and gesture: Insights from pointing. *Glossa* 4(1), 2, 1-26.
Fenlon, J., A. Schembri & K. Cormier. 2018. Modification of indicating verbs in British Sign Language: A corpus-based study. *Language* 94(1), 84-118.
Finkl, T., A. Hahne, A.D. Friederici, J. Gerber, D. Mürbe & A. Anwander. 2020. Language without speech: Segregating distinct circuits in the human brain. *Cerebral Cortex* 30(2), 812-823.
Fricke, E. 2012. *Grammatik multimodal: Wie Wörter und Gesten zusammenwirken.* Berlin: De Gruyter.
Fröhlich, M., C. Sievers, S.W. Townsend, T. Gruber & C.P. van Schaik. 2019. Multimodal communication and language origins: Integrating gestures and vocalizations. *Biological Reviews* 94(5), 1809-1829.
Fuchs, S., E. Savin, S. Solt, C. **Ebert** & M. Krifka. 2019. Antonym adjective pairs and prosodic iconicity: Evidence from letter replications in an English blogger corpus. *Linguistics Vanguard* 5(1), 1-15.
Giorgolo, G. 2010. A formal semantics for iconic spatial gestures. *Logic, Language and Meaning*, 305-314.
Ginzburg, J. & A. Lücking. To appear. I thought pointing is rude: A dialogue-semantic analysis of pointing at the addressee. To appear in *Proceedings of Sinn und Bedeutung* 25.
Goldin-Meadow, S. & M.W. Alibali. 2013. Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology* 64, 257-283.
Goldin-Meadow, S. & D. Brentari. 2017. Gesture, sign, and language: The coming of age of sign language and gesture studies. *Brain & Behavioral Science* 40, 1-60.
Goldin-Meadow, S., S. Cook & Z.A. Mitchell. 2009. Gesturing gives children new ideas about math. *Psychological Science* 20(3), 267-272.
Greenberg, G. 2018. Content and target in pictorial representation. *Ergo, an Open Access Journal of Philosophy* 5(33), 865-898.
Grosz, P., E. Kaiser & F. Pierini. 2021. Discourse anaphoricity and first-person indexicality in emoji resolution. *Proceedings of Sinn und Bedeutung* 25, 340-357. https://doi.org/10.18148/sub/2021.v25i0.94.
Hanke, T. 2016. Towards a visual sign language corpus linguistics. *Proceedings of the 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining*, 89-92.
Harrison, S. 2018. *The Impulse to Gesture: Where Language, Minds, and Bodies Intersect*. Cambridge: CUP.
Hellbernd, N. & D. Sammler. 2016. Prosody conveys speakers' intentions: Acoustic cues for speech act perception.

Journal of Memory and Language 88, 70-86.
Hobaiter, C.L., R.W. Byrne & K. Zuberbühler. 2017. Wild chimpanzees' use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology* 71, 96, 1-13.
Hohenberger, A. & H. Leuninger 2012. Production. In Pfau, R., M. Steinbach & B. Woll (eds.), *Sign Language. An International Handbook*. Berlin/Boston: De Gruyter Mouton, 711-738.
Holler, J. 2014. Experimental methods in co-speech gesture research. In Müller, C., S. Ladewig, A. Cienki, E. Fricke, J. Bressem & D. McNeill (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction.* Berlin/Boston: De Gruyter Mouton, 837-856.
Holler, J., K.H. Kendrick & S.C. Levinson. 2018. Processing language in face-to-face conversation: Questions with gestures get faster responses. *Psychonomic Bulletin & Review* 25(5), 1900-1908.
Hough, J., Y. Tian, L. de Ruiter, S. Betz, S. Kousidis, D. Schlangen & J. Ginzburg. 2016. DUEL: A multi-lingual multimodal dialogue corpus for disfluency, exclamations and laughter. *Proceedings of LREC 10*, 1784-1788.
Hübl, A., E. Maier & M. **Steinbach**. 2019. To shift or not to shift. Quotation and attraction in DGS. *Sign Language & Linguistics* 22(2), 71-209.
Hunter, J. 2019. Relating gesture to speech: reflections on the role of conditional presuppositions. *Linguistics & Philosophy* 42, 317-332.
Igualada, A., L. Bosch & P. Prieto. 2015. Language development at 18 months is related to multimodal communicative strategies at 12 months. *Infant Behavior and Development* 39, 42–52.
Janzen, T. 2012. Lexicalization and grammaticalization. In Pfau, R., M. Steinbach & B. Woll (eds.), *Sign Language. An International Handbook*. Berlin/Boston: De Gruyter Mouton, 816-841.
Joo, J., F.F. Steen & M. Turner. 2017. Red Hen Lab: Dataset and tools for multimodal human communication research. *KI – Künstliche Intelligenz* 31, 357-361.
Kaiser, E. & P. Grosz. 2021. Anaphoricity in emoji: An experimental investigation of face and non-face emoji. *Proceedings of the Linguistic Society of America* 6, 1009-1023.
Kamel Boulos, M.N., B.J. Blanchard, C. Walker, J. Montero, A. Tripathy & R. Gutierrez-Osuna. 2011. Web GIS in practice X: A Microsoft Kinect natural user interface for Google Earth navigation. *International Journal of Health Geographics* 10, 1-14.
Kaplan, D. 1989. Demonstratives. In Almog, J., J. Perry & H. Wettstein (eds.), *Themes from Kaplan*. Oxford: OUP, 565-614.
Kelly, S.D., T. McDevitt, & M. Esch. 2009. Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes* 24(2), 313-334.
Kendon, A. 1980. Gesticulation and speech: two aspects of the process of utterance. In M. R. Key (ed.), *Nonverbal Communication and Language.* The Hague: Mouton, 207-227.
Kendon, A. 1995. Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics* 23(3), 247-279.
Kersken, V., J.-C. Gómez, U. Liszkowski, A. Soldati & C. Hobaiter. 2018. A gestural repertoire of 1- to 2-year-old human children: in search of the ape gestures. *Animal Cognition* 22(4), 577-595.
Kita, S. 2009. Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes* 24(2), 145-167.
Kita, S. & A. Özyürek. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48(1), 16-32.
Kopp, S. 2010. Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors. *Speech Communication* 52(6), 587-597.
Kopp, S., P. Tepper & J. Cassell. 2004. Towards integrated microplanning of language and iconic gesture for multimodal output. *Proceedings of the 6th International Conference on Multimodal Interfaces*, 97-104.
Krause, M.A., M.A.R. Udell, D.A. Leavens & L. Skopos. 2018. Animal pointing: Changing trends and findings from 30 years of research. *Journal of Comparative Psychology* 132(3), 326-345.
Krenn, B., C. Pelachaud, H. Pirker & C. Peters. 2011. Embodied conversational characters: Representation formats for multimodal communicative behaviours. In Petta, P., C. Pelachaud & R. Cowie (eds.), *Cognitive Technologies. Emotion-Oriented Systems: The Humaine Handbook.* Heidelberg: Springer, 389-415.
Kripke, S. 1977. Speaker's reference and semantic reference. *Midwest Studies in Philosophy* 2(1), 255-276.
Krivochen, D.G. & L. Lacková. 2020. Iconicity in syntax and the architecture of linguistic theory. *Studies in Language* 44(1), 95-131.
Kushch, O., A. Igualada & P. Prieto. 2018. Prominence in speech and gesture favor second language novel word learning. *Language, Cognition and Neuroscience* 33(8), 992-1004.
Ladewig, S.H. 2020. *Integrating Gestures. The Dimension of Multimodality in Cognitive Grammar*. Berlin: De Gruyter Mouton.
Ladewig, S.H. & L. Hotze. In press. Erzählen als dynamischer multimodaler Prozess. In Zima, E. & C. Weiß (eds.), *Multimodales Erzählen. Special Issue of Linguistik Online*.
Lascarides, A. & M. Stone. 2009. A formal semantic analysis of gesture. *Journal of Semantics* 26(4), 393-449.
Leonard, M.K., B. Lucas, S. Blau, D.P. Corina & E.F. Chang. 2020. Cortical encoding of manual articulatory and linguistic features in American Sign Language. *Current Biology* 30, 1-10.

Li, Y., J. Huang, F. Tian, H.-A. Wang & G.-Z. Dai. 2019. Gesture interaction in virtual reality. *Virtual Reality & Intelligent Hardware* 1(1), 84-112.
Liddell, S. 2003. *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: CUP.
Liebal, K., C. Müller & S. Pika (eds.). 2005. Gestural communication in nonhuman and human primates. Special issue of *Gesture* 5(1-2).
Lücking, A. 2013. *Ikonische Gesten. Grundzüge einer linguistischen Theorie*. Berlin/Boston: De Gruyter.
Lücking, A., K. Bergman, F. Hahn, S. Kopp & H. Rieser. 2013. Data-based analysis of speech and gesture: The Bielefeld Speech and Gesture Alignment Corpus (SaGA) and its applications. *Journal of Multimodal User Interfaces* 7(1-2), 5-18.
Lücking, A., A. Mehler, D. Walther, M. Mauri & D. Kurfürst. 2016. Finding recurrent features of image schema gestures: The FIGURE corpus. *Proceedings of LREC 10*, 1426-1431.
Lücking, A., T. Pfeiffer & H. Rieser. 2015. Pointing and reference reconsidered. *Journal of Pragmatics* 77, 56-79.
Lüke, C. & U. Ritterfeld. 2014. The influence of iconic and arbitrary gestures on novel word learning in children with and without language delay. *Gesture* 14, 204-225.
Lüke, C., U. Ritterfeld, A. Grimminger, K.J. Rohlfing & U. Liszkowski. 2020. Integrated communication system: Gesture and language acquisition in typically developing children and children with LD and DLD. *Frontiers in Psychology* 11 (118), 1-13.
Lupyan, G. & B. Winter. 2018. Language is more abstract than you think, or, why aren't languages more iconic? *Philosophical Transactions of the Royal Society* B 373(1752), 1-11.
Macedonia, M., K. Müller & A.D. Friederici. 2011. The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping* 32(6), 982-998.
Maier, E. 2018. Quotation, demonstration, and attraction in sign language role shift. *Theoretical Linguistics* 44(3/4), 165-176.
Maier, E. 2019. Picturing words: The semantics of speech balloons. *Proceedings of the 22nd Amsterdam Colloquium*, 584-592.
Maier, E. & M. **Steinbach**. 2022. Perspective shift across modalities. *Annual Review of Linguistics* 8, 59-76.
McNeill, D. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
McNeill, D. 2005. *Gesture and Thought*. Chicago: University of Chicago Press.
Meier, R.P. 2002. Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In Meier, R.P., K. Cormier & D. Quinto-Pozos (eds.), *Modality and Structures in Signed and Spoken Languages*. Cambridge: CUP, 1-25.
Mittelberg, I. 2018. Gestures as image schemas and force gestalts: A dynamic systems approach augmented with motion-capture data analyses. *Cognitive Semiotics* 11(1), 1-21.
Mondada, L. 2007. Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers.
Morgenstern, A. & M. Guidetti (eds.). 2017. The Gesture-Sign Interface in Language Acquisition. Special issue of *Language, Interaction and Acquisition* 8(1).
Müller, C. 2014. Gestures as a medium of expression: The linguistic potential of gestures. In Müller, C., A. Cienki, E. Fricke, S.H. Ladewig, D. McNeill, S. Teßendorf & J. Bressem (eds.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*. Berlin/Boston: De Gruyter Mouton, 202-217.
Müller, C., A. Cienki, E. Fricke, S.H. Ladewig, D. McNeill, S. Teßendorf & J. Bressem (eds.). 2013/2014. *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*. 2 Volumes. Berlin/Boston: De Gruyter Mouton.
Napoli, D.J. & L. Kraus. 2017. Suggestions for a parametric typology of dance. *Leonardo* 50(5), 468-476.
Nunberg, G. 1993. Indexicality and deixis. *Linguistics and Philosophy* 16, 1-46.
Okrent, A. 2002. A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics. In Meier, R.P., K. Kormier & D. Quinto-Pozos (eds.), *Modality and Structure in Signed and Spoken Languages*. Cambridge: CUP, 175-198.
Oomen, M. 2020. *Iconicity as a Mediator between Verb Semantics and Morphosyntactic Structure. A Corpus-Based Study on Verbs in German Sign Language*. PhD Thesis. University of Amsterdam.
Orfanidou, E., B. Woll & G. Morgan 2015. *Research Methods in Sign Language Studies*. Malden, MA: Wiley Blackwell.
Özçalışkan, Ş., L.B. Adamson & N. Dimitrova. 2016. Early deictic but not other gestures predict later vocabulary in both typical development and autism. *Autism* 20(6), 754763.
Özçalışkan, Ş. & S. Goldin-Meadow. 2005. Gesture is at the cutting edge of early language development. *Cognition* 96(3), B101-B113.
Paggio, P., M. Agirrezabal, B. Jongejan & C. Navarretta. 2020. Automatic detection and classification of head movements in face-to-face communications. *Proceedings of LREC2020*, 15-21.
Parrill, F. 2010. Viewpoint in speech-gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes* 25(5), 650-668.
Patel-Grosz, P., P.G. Grosz, T. Kelkar & A.R. Jensenius. 2018. Coreference and disjoint reference in the semantics of narrative dance. *Proceedings of Sinn und Bedeutung 22*, 199-216.
Pelachaud, C. 2005. Multimodal expressive embodied conversational agents. *Proceedings of the 13th annual ACM International Conference on Multimedia*, 683-689.
Perlman, M. & N.J. Benitez. 2010. Talking fast: The use of speech rate as iconic gesture. In Perrill, F., V. Tobin & M.

Turner (eds.), *Meaning, Form, and Body*. Stanford: CSLI Publications, 245-262.
Perlman, M. & G. Lupyan. 2018. People can create iconic vocalizations to communicate various meanings to naïve listeners. *Scientific Reports* 8, 2634, 1-14.
Perniss, P. 2007. *Space and Iconicity in German Sign Language (DGS)*. PhD Thesis. Max Planck Institute for Psycholinguistics, Nijmegen.
Perniss, P. 2012. Use of sign space. In Pfau, R., M. Steinbach & B. Woll (eds.), *Sign Language. An International Handbook*. Berlin/Boston: De Gruyter Mouton, 412-431.
Perniss, P., R.L. Thompson & G. Vigliocco. 2010. Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology* 1, 1-15.
Pfau, R., M. **Steinbach** & B. Woll (eds.). 2012. *Sign Language. An International Handbook*. Berlin/Boston: De Gruyter Mouton.
Pika, S. 2008. Gestures of apes and pre-linguistic human children: Similar or different? *First Language* 28(2), 116-140.
Pouw, W., J.P. Trujillo & J.A. Dixon. 2020. The quantification of gesture–speech synchrony: A tutorial and validation of multi-modal data acquisition using device-based and video-based motion tracking. *Behavior Research Methods* 52, 723-740.
Quer, J. & M. **Steinbach**. 2019. Handling sign language data. The impact of modality. *Frontiers in Psychology* 10, 483, 1-8.
Rathmann, C. & G. Mathur. 2002. Is verb agreement the same crossmodally? In Meier, R.P., K. Cormier & D. Quinto-Pozos (eds.), *Modality and Structure in Signed and Spoken Languages*. Cambridge: CUP, 370-404.
Rieser, H. & I. Lawler. 2020. Multi-modal meaning – An empirically-founded process algebra approach. *Semantics & Pragmatics* 13(8) 1-48.
Rohlfing, K.J. 2019. Language learning from the use of gestures. In Horst, J. & J. v. Koss Torkildsen (eds.), *International Handbook of Language Acquisition*. New York: Routledge, 213-233.
de Ruiter, J.P. 2000. The production of gesture and speech. In McNeill, D. (ed.), *Language and Gesture*. Cambridge: CUP, 284-311.
Sandler, W., M. Aronoff, C. Padden & I. Meir. 2014. Language emergence. In Sindell, J., P. Kockelmann & N. Enfield (eds.), *The Cambridge Handbook of Linguistic Anthropology*. Cambridge: CUP, 250-284.
Sandler, W. & D. Lillo-Martin. 2006. *Sign Language and Linguistic Universals*. Cambridge: CUP.
Sansavini, A., A. Guarini, M. Zuccarini, J. Zong Lee, G. Faldella & J.M. Iverson. 2019. Low rates of pointing in 18mMonth-olds at risk for autism spectrum disorder and extremely preterm infants: A common index of language delay? *Frontiers in Psychology* 10, 2131, 1-12.
Schiel, F., S. Steininger & U. Türk. 2002. The SmartKom Multimodal Corpus at BAS. *Proceedings of LREC* 3, 200-206.
Schlenker, P. 2014. Iconic features. *Natural Language Semantics* 22(4), 299-356.
Schlenker, P. 2018a. Gesture projection and cosuppositions. *Linguistics & Philosophy* 41(3), 295-365.
Schlenker, P. 2018b. Visible meaning: Sign language and the foundations of semantics. *Theoretical Linguistics* 44(3-4), 123-208.
Schlenker, P. 2019. What is super semantics? *Philosophical Perspectives* 32, 365-453.
Schlenker, P. 2020. Gestural grammar. *Natural Language & Linguistic Theory* 38, 887-936.
Schlenker, P. 2021. Triggering presuppositions. *Glossa: A Journal of General Linguistics* 6(1) 1-28. DOI: https://doi.org/10.5334/ gjgl.1352.
Schmitt, R. 2005. Zur multimodalen Struktur von turn-taking. *Gesprächsforschung* 6, 17-61.
Schouwstra. M. 2012. *Semantic Structures, Communicative Strategies and the Emergence of Language*. PhD Thesis. University of Utrecht.
Senghas, A. & M. Coppola. 2011. Getting to the point. How a simple gesture became a linguistic element in Nicaraguan Signing. In Mathur, G. & D.J. Napoli (eds.), *Deaf around the World: The Impact of Language*. Oxford: OUP, 127-143.
Seyfeddinipur, M. 2004. Meta-discursive gestures from Iran: Some uses of the 'Pistolhand'. In Müller, C. and R. Posner (eds.), *The Semantics and Pragmatics of Everyday Gestures*. Berlin: Weidler, 205-216.
Stec, K. 2012. Meaningful shifts: A review of viewpoint markers in co-speech gesture and sign language. *Gesture* 12(3), 327-360.
Steinbach, M. 2021. Role shift – Theoretical perspectives. In Quer, J., R. Pfau & A. Herrmann (eds.), *Theoretical and Experimental Sign Language Research*. London: Routledge, 351-377.
Steinbach, M. & E. Onea. 2016. A DRT analysis of discourse referents and anaphora resolution in sign language. *Journal of Semantics* 33, 409-448.
Stokoe, W.C. 1960. Sign language structure: An outline of the visual communication systems of the American deaf. *Studies in Linguistics Occasional Papers* 8. Buffalo, NY: University of Buffalo, Department of Anthropology and Linguistics.
Streeck, J. 2009. *Gesturecraft. The Manu-Facture of Meaning*. Amsterdam: Benjamins.
Streeck, J., & U. Hartge. 1992. Previews: Gestures at the transition place. In Auer, P. & A. d. Luzio (eds.), *The Contextualization of Language*. Amsterdam: Benjamins, 135-157.
Taub, S.F. 2012. Iconicity and metaphor. In Pfau, R., M. Steinbach & B. Woll (eds.), *Sign Language. An International Handbook*. Berlin/Boston: De Gruyter Mouton, 388-412.
Tieu, L., R. Pasternak, P. Schlenker & E. Chemla. 2018. Co-speech gesture projection: Evidence from truth-value

judgment and picture selection tasks. *Glossa* 2(1), 102, 1-27.
Tieu, L., P. Schlenker & E. Chemla. 2019. Linguistic inferences without words. *Proceedings of the National Academy of Science (PNAS)* 116(20), 9796-9801.
Tomasello, M. 2006. Why don't apes point? In Enfield, N.J. & S.C. Levinson (eds.), *Roots of Human Sociality: Culture, Cognition and Interaction*. Oxford: Berg, 506-524.
Tomasello, M. 2008. *Origins of human communication*. Cambridge, MA: MIT Press.
van Berkel-van Hoof, L., D. Hermans, H. Knoors & L. Verhoeven. 2019. Effects of signs on word learning by children with developmental language disorder. *Journal of Speech, Language, and Hearing Research* 62, 1798-1812.
van Herreweghem, M. & M. Vermeerbergen. 2012. Handling sign langauge data. In Pfau, R., M. Steinbach & B. Woll (eds.), *Sign Language. An International Handbook*. Berlin/Boston: De Gruyter Mouton, 1023-1045.
van Loon, E., R. Pfau & M. **Steinbach**. 2014. The grammaticalization of gestures in sign languages. In **Müller**, C., S. Ladewig, A. Cienki, E. Fricke, J. Bressem & D. McNeill (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction.* Berlin/Boston: De Gruyter Mouton, 2133-2149.
Vigliocco, G., P. Perniss & D. Vinson. 2014. Language as a multimodal phenomenon: Implications for learning, processing and evolution. *Philosophical Transactions of the Royal Society* B 369, 1651, 1-7.
Vilà-Giménez, I. & P. Prieto. 2020. Encouraging kids to beat: Children's beat gesture production boosts their narrative performance. *Developmental Science* 23(6), 1-14.
Vogel, R. & M. Huth. 2020. Modusschnittstellen in mathematischen Lernprozessen. Handlungen am Material und Gesten als diagrammatische Tätigkeit. In Kadunz, G. (ed.). *Zeichen und Sprache im Mathematikunterricht. Semiotik in Theorie und Praxis*. Berlin: Springer, 215-253.
Vogt, S. & C. Kauschke. 2017. Observing iconic gestures enhances word learning in typically developing children and children with specific language impairment. *Journal of Child Language* 44, 1458-1484.
Wagner, P., Z. Malisz & S. Kopp. 2014. Gesture and speech in interaction: An overview. *Speech Communication* 57, 209-232.
Wei, S.-E., V. Ramakrishna, T. Kanade & Y. Sheikh. 2016. Convolutional pose machines. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4724-4732.
Weisberg, J., A.L. Hubbard & K. Emmorey. 2016. Multisensory semantic integration of spontaneously produced representational co-speech gestures: An fMRI study. *Language, Cognition and Neuroscience* 32(2), 158-174.
Weismer, S.E. & L.J. Hesketh. 1993. The influence of prosodic and gestural cues on novel word acquisition by children with specific language impairment. *Journal of Speech, Language, and Hearing Research* 36(5), 1013-1025.
Wilbur, R. 2013. The point of agreement. Changing how we think about sign language, gesture, and agreement. *Sign Language & Linguistics* 16(2), 221-258.
Wolfe, R., E. Efthimiou, J. Glauert, T. Hanke, J. Schnepp & J. McDonald (eds.). 2016. Recent advances in sign language translation and avatar technology. Special issue of *Universal Access in the Information Society* 15(4).