Parallel Session 1 – Tuesday, 8th 10:20 am – 12:00 pm

Teachers' gestures in instruction: Panel

Abstract word or concrete word: How the gestures of future language teachers differ

Marion Tellier, Gale Stam

Speakers adapt their speech to their interlocutors, and when they talk to non-native speakers, they tend to engage in foreigner talk (Ferguson, 1975). They use more basic vocabulary, shorter sentences, and present tense. They articulate more, speak more slowly, talk more loudly, and use gestures. The speech adjustments that speakers make in addressing non-native speakers (Wesche & Ready, 1985), and their effectiveness in facilitating acquisition (Long, 1980) have been explored within second language acquisition research. In addition, Adams (1998) examined the adjustments that native speakers make in their gestures when addressing non-native interlocutors. He found that only deictic gestures were significantly more numerous in the non-native condition and that the speakers produced about the same amount of metaphorics and emblems with both native and non-native interlocutors.

The question arises as to what future foreign language teachers do. Do they use the same types of gestures in addressing native and non-native speakers of a language. We know that foreign language teachers tend to gesture a lot in a classroom, (Tellier, 2008; Sime, 2008), that these "teaching gestures" capture attention and make the lesson more dynamic, that they support comprehension, and that they are relevant for learners' memorization processes. There is also evidence that future language teachers adjust their gestures to whether they are explaining words to native or non-native speakers (Tellier & Stam, 2012). However, we don't know whether the type of words future language teachers are explaining, for example concrete vs. abstract, has an effect on the types of gestures they use. In other words, do future teachers adjust their gestures to the types of words they are explaining to native and non-native speakers'.

10 future French teachers in training session were asked to explain 12 words to two different partners: one native speaker of French and one non-native (with an intermediate level of French, B1/B2 of the CEFR) in order to have them guess the words. Some of them were concrete (e.g., grimper "to climb" , rapidement "quickly" , trottoir "pavement,"); others were abstract (jalousie "jealousy", approximativement "approximately" , se souvenir "to remember" and fiére "proud"). It was hypothesized that regardless of the nature of the word, there would be more iconics in the NN condition.

In this talk we will discuss how the type of word being explained affected the gestures of the trainees. We looked at type, size and duration of gestures in both conditions (native vs. non- native) for each type of word (abstract vs. concrete). We found more iconics were produced with concrete words but not abstract words in the NN condition. With abstract words, more metaphorics and emblems were produced. It's not the nature of the interlocutor alone that affects the type of gestures used but also the type of words explained. Still, preliminary results show that gesture rate and gesture duration are affected by the interlocutor: in the non-native condition, gesture rate is higher and gestures last longer than in the native condition.

How instructors connect ideas using speech and gesture: Evidence from statistics lessons

Martha Alibali, Amelia Yeo, Elise Lockwood, Noelle Crooks, Mitchell Nathan

What factors affect whether students grasp connections among ideas'. Recent research has targeted teachers' instructional communication as one important factor. In mathematics and statistics instruction, teachers often highlight relationships among ideas. In doing so, teachers typically use speech, and along with that speech they often produce gestures.

The primary aim of this research was to characterize how statistics instructors connect ideas in introductory statistics courses. In particular, we focus on instruction about confidence intervals (CIs), an important topic in statistical reasoning. We investigated whether instructors typically used multiple modalities (e.g., speech, gesture or writing/drawing) to express linked ideas, or whether they sometimes expressed linked ideas in a single modality. In light of past research documenting multi-modal linking in middle-school teachers'mathematics lessons (Authors blinded, Date), we hypothesized that statistics instructors would often express linked ideas multi-modally.

A secondary aim of this research was to investigate the range of ways instructors use gestures to connect ideas. We focused on two dimensions that seemed potentially important: the types of gestures (in particular, pointing vs. depictive gestures), and the timing of the gestures.

We videotaped four university-level instructors as they taught about confidence intervals in introductory statistics courses. We transcribed each lesson and identified linking episodes, defined as segment of discourse in which instructors explicitly made connections between two or more different ideas. We then coded the modalities instructors used to express each of the linked ideas. We categorized gestures into three categories, using an adaptation of McNeill's (1992) coding scheme: points, depictive gestures, or writing gestures (i.e., when instructors made marks on the board while speaking). For links that included two or more ideas expressed in

gestures, we also coded whether the gestures that expressed the linked ideas were sequential (i.e., first one idea and then another) or simultaneous (i.e., more than one idea at a the same moment).

On average, instructors expressed all of the linked ideas multi-modally in 62% of links. Although this represents a majority of all links, this value is much lower than the comparable value of 90% for the middle-school lessons (Authors blinded, Date). There was substantial variation across instructors in the proportion of links expressed multi-modally (range 40% to 100%).

Instructors used a variety of different types of gestures to express linked ideas. Depictive gestures were most common overall (58% of the 253 coded gestures), followed by points (32%) and writing gestures (10%). Instructors often mixed gesture types within the same link.

When instructors used gestures to communicate links, they typically gestured to linked ideas sequentially. In only 2 cases (out of 36 cases in which multiple linked ideas were expressed in gesture) did instructors express linked ideas in simultaneous gestures.

In sum, gesture is pervasive in instruction on CIs, and that it is used to convey important information about connections among ideas. This work paves the way for future studies that will test the implications of variations in instructors' linking for students' statistics learning.

Teachers' attitudes and beliefs about the utility of gestures in classroom learning and instruction

Mitchell Nathan, Rebecca Boncoddo

Increasingly, scholars are investigating the nature of gestures as they occur during classroom instruction, including the types of gestures teachers produce, the circumstances that appear to elicit them, and the influence they have on student understanding and learning (e.g., Alibali et al., 2013a, 2013b; Roth, 2001). However, little is known about teachers' views about gestures during instruction. This study reports on teachers' attitudes and beliefs about gestures, using a survey developed for this purpose.

Teachers from the USA (N = 192; 39 males) were recruited using email lists from several school districts and research collaborators, and offered a gift card to a major online vendor for compensation. The sample included pre-school (n = 2), primary (grades K-5, n = 65), middle (grades 6V8, n = 51), and secondary (grades 9V12, n = 74) educators. Reliability (Cronbach's alpha) was established for the survey with 181 participants from a prior cohort of teacher education students. Planned comparisons showed no differences due to teacher gender and no grade-level differences. Survey items used 5-point Likert scales.

We found that teachers reported gestures as mildly beneficial for student learning (Mean = 3.59; alpha = .844) and as not distracting for students (Mean = 2.13; alpha = .795). There was a significant negative correlation between these two scales, r = -.24, suggesting that teachers believe gestures contribute to effective instruction without distracting learners.

The distinction between gesture-speech matches and mismatches (i.e., redundant vs. complementary gestures) is important in gesture studies (Church & Goldin-Meadow, 1986; Singer & Goldin-Meadow, 2005). Teachers strongly agreed that gesture-speech matches contribute to student learning as well as their own learning (Mean = 3.93; alpha = .596). Yet teachers also agreed that gestures that are complementary to concurrent speech contribute to students' understanding of instruction (Mean = 3.79; alpha = .566).

We also asked teachers to consider the perceived causes of the efficacy of gesture in instruction (5 items; alpha = .871). Teachers strongly held (Mean = 4.17) the view that when gestures were effective it was because they: (1) helped make clear connections between ideas or representations; (2) made the interaction more engaging for the learner; (3) helped make abstract ideas seem more concrete; (4) helped the teacher to formulate their speech; and (5) helped to direct the learners' attention to the curriculum material. Teachers did not particularly agree that they gestured to serve as an aid to formulate their thoughts into words, or for the sake of being theatrical (6 items; alpha = .80; Mean = 3.3). Rather, teachers more strongly agreed (Mean = 3.78) that they used gestures to clarify something, act something out, connect ideas, connect representations to one another, and to help make abstract ideas more concrete for students.

Avatar-based research on gesture in instruction: Opportunities, challenges, and solutions

Jian Cui, Meng-Lin Wu, Suren Deepak Rajasekaran, Nicoletta Adamo-Villani, Voicu Popescu

Research on gesture in instruction has traditionally relied on video stimuli. Video stimuli are by definition photorealistic and viewers have no trouble associating the instructor in the video to a real life instructor. However, recording video stimuli relies on an instructor to execute a script, which is challenging. First, the instructor has to learn the script to perform it precisely as written. Second, the instructor has to perform multiple versions of the script, one for each of the conditions of the study, while maintaining all secondary parameters constant (e.g. speech intensity, energy, enthusiasm, head motion, body pose, and secondary arm and hand motions). Consequently, creating video stimuli can be a long and tedious process, and video stimuli can appear constrained and unnatural.

Computer animation characters serving as instructor avatars have the potential to bypass many of these challenges. Avatars have perfect memory, infinite energy, and endless enthusiasm. Avatars enable exact control of the stimuli parameters. For example, all conditions can use the exact same audio and the exact same secondary motion. Avatar stimuli are potentially simpler to make, enabling complex studies, with many conditions, that can discern the individual contribution and the interplay of many gesture variables. However, with

current software tools, creating avatar-based stimuli requires artistic talent and technical expertise. We report on our work to overcome this challenge and to make computer avatars an accessible yet powerful platform for research on gesture in instruction.

Education researchers will not and should not create their own avatars; instead, a database of avatars should be built and shared. We have built two avatars, Julie and Jason. Algorithmic animation, like pointing to a symbol on a white board, is computed and executed on demand. We have developed inverse kinematics algorithms for pointing, circling, underlining, and tapping at any location in 3-D space. More complex animation, like a balance gesture used for learning mathematical equivalence, has to be constructed beforehand by an animator and invoked when needed. Our database contains instructor poses, embodied cognition and charisma gestures, and facial expressions. The avatar also has to speak. Synthesizing speech from text results in robotic speech. Instead we record speech to which we lip-sync the avatar automatically using pre-animated set of visemes.

The last required ingredient is a convenient interface for the education researcher to assemble the stimuli needed in the studies. We are developing a scripting language that formalizes the description of the environment (e.g. what is written on the whiteboard), the avatar commands (e.g. speak, gesture, assume pose), and the synchronization between speech and gesture. The script can be created through a text editor, or through a graphical user interface that lowers the scripting language learning curve.

Our avatars are used in gesture studies that research equivalence relations, linking ideas, and instructor charisma in mathematical instruction. The avatars allow building high-quality stimuli quickly and with perfect control of secondary parameters. These studies show that instructor avatars are promising technological infrastructure for research on gesture in instruction.

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A130016, and by the National Science Foundation, through Grant 1217215. The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences, of the U.S. Department of Education, or of the National Science Foundation.

Sign & gesture 1: Talks

Natural sign in Nepal

E. Mara Green

This paper draws on linguistic anthropological research in Nepal to examine semi-conventional signed practices known by Nepali Sign Language (NSL) users as "natural sign." The meta-linguistic category natural sign encompasses overlapping communicative repertoires used by (1) deaf people who have not acquired NSL, (2) hearing people when communicating with deaf persons, and (3) NSL signers when communicating with either of the former. Focusing on natural sign's linguistic properties as well as its acquisition and use, I will argue that natural sign manifests similarities with co-speech gesture, home sign, and (village) sign languages but constitutes a distinct category of manual-visual communication.

As I will demonstrate with video data, natural sign consists of a small number of conventionalized lexical items that vary across social/geographical spaces; their combination and use are governed by basic syntactic rules and pragmatic strategies. Like co-speech gesture (McNeill 2000; Kendon 2004), natural sign sometimes accompanies speech, whether the signer is deaf, hard of hearing, or hearing. In many cases, however, natural sign is produced with minimal or no vocal articulations or only with mouthing.

Natural sign is also different from co-speech gesture in that the former functions as a primary mode of communication between hearing and deaf people as well as among deaf people. Indeed as the only mode used by deaf people who have not acquired NSL, natural sign is similar to home sign (e.g. Morford and Kegl 2000). While scholarly work frames home sign as the de novo creation of an individual deaf person (e.g. Goldin-Meadow 2003), however, NSL signers describe how as young children (before encountering NSL), they learned natural sign in interactions with hearing relatives and neighbors. This account is reminiscent of village sign languages (e.g. Nyst 2007, Sandler et al 2011). Natural sign, however, is less systematized than village sign languages and there is significant variation in how well individuals can produce and understand natural sign. Relatedly, NSL signers say that it is easy to communicate in natural sign but that it also imposes limits. In the words of one young man, It's a smaller thing. At the same time, natural sign is a "larger thing" in that a wide range of people uses it V from urban-based NSL signers to deaf villagers, from the hearing relatives of deaf people to bus-drivers and shop-owners. Village sign languages, in contrast, are localized in particular places where there is a high incidence of familial, multigenerational deafness.

In sum, my research presents a distinct category of manual-visual communication that is durable and socially available to both deaf and hearing people. The degree of conventionalization across different social and geographic spaces suggests that natural sign is distinct from co-speech gesture and home sign; its variability and limits suggest that it is distinct from sign language.

Movement and Handshape in Gesture and Sign Language: A Cross Linguistic, Cross Cultural Study

Laura Horton, Diane Brentari

In this study we consider the distribution of information encoded in movement and handshape in sign language and gesture and posit that this distribution is affected by culture, in the form of "gestural context," as well as the presence or

absence of a linguistic system. We ask first whether movement is used consistently by both gesturers and signers to provide information about the plurality of an event and second, whether the information about agency in movement is isomorphic with the information about agency in handshape.

In sign languages, handshape conveys information about the presence or absence of an agent in an event1, while movement has the capacity to communicate information about agency as well as plurality2. It is not clear that gesturers use handshape contrastively as signers do, and preliminary evidence indicates that the handshape strategy used by signers may be more salient for gesturers from richer co-speech gestural contexts. This study builds on this body of work with experimental evidence that signers and gesturers use movement to encode information about plurality, but that different sign languages distribute information about agency and plurality differently in the components of movement and handshape.

Methods: This study compared productions by signers of American and Italian Sign Languages, and hearing speakers of English and Italian. Hearing participants used only gesture to communicate. All participants (16 total) described an equal number of vignettes of each of the following types: a single object, multiple objects, a hand placing a single object and a hand placing multiple objects. 762 productions were coded for movement type and handshape type. Movement types included contact single, path single, contact multiple and path multiple. These types were characterized by a combination of the following features: movement shape (straight or arc), axis (vertical or mid-sagittal) and repetition (single or multiple). The handshapes used in each production were coded according to how the participant's handshape described the object. The handshape could mimic the shape of the object or the way a hand would manipulate the object.

In descriptions of events that did not include an agent, all groups used movement to indicate plurality and both movement and handshape to encode information about agency. To describe conditions with an agent, movement was the most salient strategy to provide information about both plurality and agency for gesturers. Culturally, the agent information in handshape was more consistent with the agent information in movement for Italian subjects (LIS signers and Italian gesturers). For signers (ASL and LIS), movement often provided information about plurality but was less reliable as an indicator of the agentive status of the elicitation event. This underscores the role of a fully developed system of meaning-form relationships in sign languages which permits signers to use a form in which one component of the construction (movement) describes one feature of the event while a separate component (handshape) reliably describes other features of the event (agency).

Organizing a lexicon with patterned iconicity

So-One Hwang, Sharon Seegers, Ryan Lepic, Elizabeth Hodgdon, Nozomi Tomita, Deniz Ilkbasaran

Perniss, Thompson, and Vigliocco (2010) suggest that iconicity in language serves to reduce the conceptual gap between linguistic form and human experience. Here, we investigate pervasive iconicity in sign languages and explore its communicative and cognitive underpinnings. We distinguish between two distinct types of iconicity in sign language lexicons: 1) of individual signs and 2) strategies that span groups of lexical signs sharing semantic properties. Focusing on the latter, we compare spontaneously created silent gestures in non-signing hearing individuals with lexical signs produced by signers in 5 different sign languages. While individual gestures and signs can vary widely in how they reflect the human experience (representing the trunk of a tree, or the bough of a tree), we find that strategies for gestures or lexical signs within semantic categories are remarkably similar across gesturers and signers. Further, as we compare how strategies are used across sign languages, we uncover subtle differences between them. Our study proposes a basis for distinguishing between general communicative resources underlying sign and gesture from grammatical organization in sign language lexicons.

Using pictures of 90 common objects across 3 semantic categories: 30 tools, 30 animals, and 30 fruits & vegetables, we elicited signs from native deaf signers of the following languages: American Sign Language (ASL), Japanese Sign Language (JSL), Israeli Sign Language (ISL), Al-Sayyid Bedouin Sign Language (ABSL), and Central Taurus Sign Language (CTSL). ASL, JSL, and ISL are urban sign languages ranging in age from young to established, and ABSL and CTSL are young village sign languages of Israel and Turkey, respectively. We also asked 12 hearing non-signers in the U.S. to produce silent gestures for the same set of pictures. We coded participants' use of their own body in their responses, and 1) whether the body (head and torso) is used to represent the object or animal in the photograph, or 2) if the body is a human agent interacting with the object or animal in the picture. (There are no human agents featured in any of the pictures.) We also coded participants' hands in their responses, whether: 1) they were used to represent any aspect of the object in the picture (such as the thin, long shape of a toothbrush) or 2) they were human hands interacting with the object in the picture (such as holding and using a toothbrush).

We found that signers and gesturers very consistently use their body and hands (at nearly 100%) to show human action involving tools when identifying pictures of such objects, though no humans or actions were shown in any of the pictures. While gesturers are more likely to show holding or grasping an imaginary tool, sign languages are more likely to use handshapes representing the shape or dimension of the tool, with the two younger village sign languages (ABSL and CTSL) showing higher use of these forms. For both gesturers and signers, pictures of fruits & vegetables are most likely to elicit forms where the hands are used to show shape and size without human action. However, the same pictures often elicit multiple responses in gesturers who use different

strategies, such as showing the shape of a tomato then how to prepare tomatoes for eating using a knife. Finally, pictures of animals are likely to elicit whole body forms where the participant's body stands in for the body of the non-human animate.

Metaphoric iconicity in signed and spoken languages

Defu Yap, Laura Staum Casasanto, Daniel Casasanto

Since Saussure, the idea that the forms of words are arbitrarily related to their meanings has been widely accepted. Yet, implicit metaphorical mappings may provide opportunities for iconicity throughout the lexicon. We hypothesized that vertical spatial metaphors for emotional valence are manifested in language through space in signed languages and through the spatialized dimension of pitch in spoken languages. In Experiment 1, we analyzed the directions of the hand motions constituting words in three signed languages, and related them to the valence of their English translation equivalents. The vertical direction of signs predicted their valences. On average, signs with upward movements were the most positive in valence, and signs with downward movements the most negative. Signs with non-vertical movements were intermediate in valence. Experiment 2 extended this type of analysis to a tonal language, Mandarin Chinese. The pitch contours of Chinese words predicted the valence of their English translation equivalents. These results demonstrate a previously unrecognized source of non-arbitrariness in language, revealing that implicit space-valence metaphors are encoded in the forms of words in both signed and spoken languages.

Interaction 1: Talks

Interactive behavioral alignment during joint remembering

Lucas Bietti, Alan Cienki, Kasper Kok

Individuals engaged in social interactions often align their communicative resources in order to achieve shared goals. Dale and colleagues claim that alignment is one way of reducing the cognitive load of interlocutors, thus condensing the complexity of the interaction (Dale et al. 2014, p. 69). Alignment has been widely investigated in relation to speech (e.g. Allen et al. 2011; Pickering & Garrod 2004). In connection to co-verbal gesture and non-verbal resources, it has been demonstrated that during naturalistic conversations people tend to mimic co-speech manual gestures (e.g. Bergmann & Kopp, 2012), synchronize eye-gaze direction (e.g. Richardson et al. 2007); and coordinate postural sways (e.g. Shockley et al. 2007).

Although most of these studies have been focused on analyzing the temporal dynamics of single behaviors (e.g. co-speech manual gesture or eye-gaze direction), new theoretical perspectives suggest that different behavioral resources constitute a holistic web of mutual inter-animations (Dale et. al., 2014; Louwerse et. al., 2012).

Various explanations for alignment have been proposed in relation to speech. While some claim that it is highly automatic and caused by priming effects (e.g. Pickering & Garrod, 2009), others stress its conscious aspects (e.g. Brennan et al., 2010).

These perspectives have led some researchers to conduct studies on linguistic alignment in relation to transactive memory, that is, in cases where the memory of the group goes beyond the sum of the individual memories of each team member (e.g. Tollefsen et al. 2013). However, this new line of inquiry has not examined how instances of alignment of co-verbal gesture and non-verbal behaviors during social interaction provide the interactional architecture for collaborative remembering in small groups (Bietti et al., 2013; Cienki et al., 2014).

In this talk, we investigate the roles that the interactive alignment of manual gesture, postural sway and eye-gaze play in small groups engaged in collaborative remembering. All participants were native Spanish speakers and video recordings were collected in real-world environments (participants' homes in Buenos Aires). Qualitative and quantitative analyses of a video corpus demonstrate how the alignment of co-speech gesture, postural sway, and eye gaze has different interactional dynamics while interactionally fostering collaborative remembering among small groups. Gestural alignment evoked shared attention and enhanced agreement. Postural alignment coincided with mutual engagement and shared thinking whereas coordination of gaze was associated with directing of others' attention. Moreover, these analyses show whether and how the differences in these behaviors' roles in joint remembering are reflected in the temporal dynamics of the alignment patterns observed in our data. Afterwards, in order to examine whether the instances of simultaneous and sequential alignment during joint remembering found in our data set can be associated to either priming effects (automatic) or conscious aspects of the interactions (task-mediated), we report preliminary results of an agent-based computer simulation. Finally, we discuss the potential of combined qualitative-quantitative approaches for illuminating the interplay of verbal and bodily coordination during contexts such as interactive memory construction.

Multimodal action-formation in "Passing-by interactions' among clinical staff

Esther González-Martínez, Kim Lê Van, Adrian Bangerter

"Passing-by interactions" constitute a specific form of unscheduled, fleeting, on-the-move encounter in which two or more participants, following opposite and close parallel trajectories, start talking to each other as they approach, pass by and/or move away from each other without stop walking. The paper is based on detailed multimodal analysis of a collec-

tion of videotaped passing-by work interactions among staff members in the corridors and other interstitial spaces of a hospital's outpatient clinic. The paper will specifically focus on the issue of multimodal action-formation: how talk, bodily conduct and the resources of the environment are articulated to produce conformations recognizable as particular actions (Schegloff, 2007, p. xiv). We will argue that, in our excerpts, talk is carefully designed and articulated to bodily conduct to be recognizable as accomplishing particular actions projecting and requesting only minimal response, which contributes to the fleeting dimension of the encounters.

Identifying "civil inattention rituals" , "passing greetings" , or "terminal squirms" , Goffman (1963, 1971) spotted some of the slightest forms of unfocused and focused interaction both in ordinary and institutional settings, including mental health hospitals. Analytically as well as methodologically, our study is part of the "spatial and mobility turn" in studies of social practices and interaction (McIlvenny et al., 2009). Investigations of on-the-move interactants establishing interactional spaces in other than static face-to-face arrangements (De Stefani, Mondada, 2010) are especially relevant to us. We bring into play multimodal analysis (Streeck et al., 2011) to study the articulation of talk, gesture, gaze, movement and other bodily conduct in the interactive production of social action.

The paper is based on an ongoing three-year video-based field study in the outpatient clinic of a hospital in the French-speaking part of Switzerland. We have collected a corpus of three hundred thirty-six hours of video using a set-up composed of four cameras and eight wireless microphones placed in the corridors of the clinic, recording simultaneously over the course of seven consecutive days, twelve hours per day. So far, the collection of passing-by encounters consists of 32 instances.

Gaze and the organization of turn-taking in triadic face-to-face interaction

Judith Holler, Kobin Kendrick

The primordial site of conversation is face-to-face social interaction where participants make use of visual modalities, as well as talk, in the coordination of collaborative action (Clark, 1996). This observation leads to a fundamental question: what is the place of multimodal resources such as these in the organisation of turn-taking for conversation'o To answer this question, we collected a corpus of both dyadic and triadic face-to-face interactions between adult native English speakers, with the aim to build on existing observations of the use of visual bodily modalities in conversation (e.g., Duncan, 1972; Goodwin, 1981; Kendon, 1967; Lerner 2003; Mondada 2007; Oloff, 2013; Rossano, 2012; Sacks & Schegloff, 2002; Schegloff, 1998). The corpus retains much of the spontaneity and naturalness of everyday talk while combining it with state-of-the-art technology to allow for exact, detailed analyses of verbal and visual conversational behaviours. Each participant (1) was filmed by three high definition video cameras (providing a frontal plus two lateral views) allowing for fine-grained, frame-by-frame analyses of bodily conduct, as well as the precise measurement of how individual bodily behaviours are timed with respect to each other, and with respect to speech; (2) wore a head-mounted microphone providing high quality recordings of the audio signal suitable for determining on- and off-sets of speaking turns, as well as inter-turn gaps, with high precision, (3) wore head-mounted eye-tracking glasses to monitor eye movements and fixations overlaid onto a video recording of the visual scene the participant was viewing at any given moment (including the other [two] participant[s] and the surroundings in which the conversation took place). The HD video recordings of body behaviour, the eye-tracking video recordings, and the audio recordings from all 2/3 participants engaged in each conversation were then integrated within a single software application (ELAN) for synchronised playback and analysis. All data have been transcribed, coded for co-speech gestures and gaze fixations on a frame-by-frame basis. The large amount of data obtained from this corpus is currently being analysed both qualitatively and quantitatively. The project aims to shed light on the cognitive puzzle that turn-taking presents us with (Levinson, 2013); interlocutors are confronted with the challenge of comprehending an on-going turn while, at the same time, planning a response and estimating when the current speaker's talk will end in order to time their contribution as precisely as possible (the average gap between turns is a mere 200ms). The results from this project provide insight into the process of turn projection as evidenced by participants' gaze behaviour with a focus on the role different bodily cues play in this context. Our findings so far show that co-speech gestures may play an important role in this process by guiding the projection of upcoming turn boundaries and next actions. In all, this project elucidates the role of multi-modality in the organisation of turns at talk and in the cognitive processes that underlie this organisation.

Deaf-hearing gestural interaction in Mumbai: An ethnography of communication

Annelies Kusters

Gestural communication is the main means of communication between deaf and hearing people in the majority of countries in the world, such as in India. The aim of this study is to understand the potential and limits of gesture use in language contact situations between deaf and hearing people who do not have fluency in a shared language (mode). In the sociolinguistically diverse environment of Mumbai, where (co-speech) gesture is widely and effectively used among speakers of different languages, the study considers two related issues with regard to communicative repertoire: how fluent deaf signers use gestures (conventionalised and spontaneous) to communicate with hearing non-signers; and how hearing speakers use gestures to communicate with deaf non-

speakers.

In particular, the discourse range of gestural communication as well as its limitations and potential are investigated. The role of speech and writing in gestural communication is analysed, as is the role of the location of the interactions (i.e., the immediate physical/spatial environment). Hearing and deaf participants' own perceptions of the relative ease of communicating on various topics in a range of situations are investigated. Relationships between the way gesture is used and the place where the interaction happens, as well as the underlying perspectives regarding gestures, sign language and deaf people, are identified.

To this end, gestural interactions in public and parochial spaces (such as markets, shops, streets, food joints, public transport, parks) between strangers, acquaintances or neighbours in Mumbai are observed and video-recorded, and interviews are conducted with both deaf and hearing individual participants to find out more about their views on gesture. The recordings provide data for analysis but also material for a film documentary, which will serve as a basis for further exploration in a second round of data collection in which the documentary will be presented for group discussion.

This research contributes to understanding the multilingual and multimodal repertoire that speakers could use to achieve communication across diverse communities when attempting to reach mutual understanding. Also, the attention for the phenomenological experience of performing or viewing gestures is an important intervention in Gesture Studies as this is a neglected dimension of gesture-use.

Experimental 1: Talks

Creative with the truth: Individual differences in nonverbal indicators of deception

Daniel Gurney

Previous research has attempted to identify nonverbal factors that reveal deception, though with limited success. It is only recently that research has revealed how the study of gestures can be useful in making veracity judgements, and studies have highlighted that people tend to gesture differently when recalling a true or fabricated story (Caso, Maricchiolo, Bonaiuto, Vrij, & Mann, 2006; Hillman, Vrij, & Mann, 2011). The difference in gesture behaviour could be explained by the additional cognitive demands placed on individuals when being dishonest (Vrij, Fisher, Mann & Leal, 2006), which gestures can serve to reduce (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Wagner, Nusbaum, & Goldin-Meadow, 2004). However, the role of representational gestures in deceptive story telling is still unclear and warrants further investigation.

The present study explores the relationship between veracity and representational gestures further by introducing praxis as an additional variable: Representational (iconic) gestures are more prevalent when the speaker describes praxic content (items which require the hands to function) and, while gestures accompanying non-praxic content appear to be adaptable to communicative situations, gestures accompanying praxic content are more difficult to suppress due to the high intra-personal function they serve in recalling information (Pine, Gurney & Fletcher, 2010).

The first experiment reported here studied the gesture behaviour of individuals that told both a true and false story and introduced praxis as a factor to encourage or dissuade individuals from producing representational gestures. Using a 2x2 within-subjects design, 40 participants (age M = 21.00, SD = 2.92) told both a true and false story with both praxic and non-praxic content. The results considered representational and beat gesture rates across the four stories for each of the participants and a main effect was confirmed for veracity and praxis. In addition, differences in gesture behaviour across the true and false stories became more pronounced when the stories included praxic content, and an interaction between the variables was present. These results confirm that gesture behaviour can vary according to story veracity, but add that this observation is mediated by the level of praxic content in the story.

To investigate the extent to which individual creativity mediated this effect, a second experiment measured the creativity of speakers (using the Kaufman Domains of Creativity Scale, KDOCS, Kaufman, 2012) when telling stories that varied in veracity and praxis. Individual creativity was found to be positively correlated to the number of representational gestures produced and, subsequently, creative individuals produced more representational gestures when lying than when telling the truth.

These results together further our understanding of how gesture behaviour varies with story veracity and individual creativity, but also provide further insight into the mechanisms underpinning the gesture differences observed by previous research. The implications of these findings and their reliability as a tool in deception detection will be discussed further.

Perception and execution of action are subserved by a left-hemispheric system for praxis

Ingo Helmich, Hedda Lausberg

Studies of the human faculty to match the observation of action with its execution have provided divergent results regarding its hemispheric underpinnings. Representational gestures that depict action represent a direct link between perception, motor simulation and execution. We therefore investigated in the present study whether the perception and gestural depiction of action are grounded in the same hemispheric system. Thus, we analyzed gestural demonstrations without speech in response to tachistoscopically presented everyday life action scenes. Videotaped hand movements of ten right-handed subjects were analyzed double-blind by two independent blind

raters employing the Neuropsychological Hand Movement Coding System. The results showed that motor output is facilitated when action is presented in the right visual hemifield, respectively the left hemisphere and this is the case for both unimanual and bimanual hand movements. Hand movements constructed with a preparation phase, a complex motion phase, and a retraction phase were executed with a right-hand preference, whereas movements without a complex phase structure preferred the left hand. When looking at hand movements on the functional level, i.e. gestures, representational left hand gestures increased when the action was perceived in the left visual hemifield. We conclude that the human faculty of matching the observation and execution of action is based within a left-hemispheric lateralized system. Moreover, a right-hand preference of hand movements seems to be grounded in conceptual thinking processes that are closely related to a left-hemispheric system of praxis although representational gestures can adapt to the semantic content of reference.

When do gestures reduce spatial working memory demands? The importance of mental simulation

Autumn Hostetter, Aidan Brawn

The production of representational gestures has been shown to alleviate working memory demands (e.g., Goldin-Meadow, Nussbaum, Kelly, & Wagner, 2001). Such findings are not intuitive, as it would seem that the act of planning and producing a gesture should actually require cognitive resources on the part of the speaker, not alleviate them. Why does doing something extra (e.g., planning and producing a meaningful gesture) actually reduce cognitive demands rather than increase them.

A possible explanation is presented by the Gesture as Simulated Action (GSA) framework (Hostetter & Alibali, 2008). Under this view, gestures occur as outward manifestations of the mental simulations that are involved in thinking about perceptual and motor events. As such, when a speaker is simulating a perceptual or motor event, it is actually easier to gesture about that event than to refrain from gesturing about that event. In this study, we tested the claim that gestures relieve spatial working memory demands, but only when the information being described with gesture has been learned in a visuo-spatial format.

Twenty-nine participants described short events from an animated cartoon after either reading a description of the event twice (verbal condition) or after reading the description once and then watching the event depicted in the animated cartoon (spatial condition). Participants were instructed to either gesture or not gesture with their description of each event. In addition, participants were asked to engage in a secondary task while they were giving their description; they were shown 5 digits in a grid before reading each story and were asked to recall the digits in the correct locations after giving each description. We found that participants had significantly better memory for the locations of the digits when they had gestured with their description than when they had not gestured, but only when they had seen the animated cartoon depicting the events. When participants had only read about the events, producing a gesture about the events did not lead to better memory on the secondary task compared to not producing a gesture. These results are predicted by the GSA framework and suggest that gestures are primarily beneficial to the larger cognitive system when they accompany speech that is based on spatial and motoric representations.

Perceptual salience and gesture production

Amelia Yeo, Martha Alibali

How might the perceptual salience of an observed action influence gesture production'o According to the Gesture as Simulated Action framework (Hostetter & Alibali, 2008), gesture production depends on the strength of the speaker's action simulation. Once the activation of an action simulation is beyond threshold, the speaker will produce a gesture. When people observe actions that are highly salient, their representations of those actions should be highly activated, and they should therefore produce gestures.

Experiment 1 manipulated the salience of manner in motion events by varying whether a goal was present. When a goal is present, the manner of motion is less salient (e.g., Gergely, Bekkering & Kiraly, 2002; Carpenter, Call & Tomasello, 2005). Fifteen native English speakers were presented with short clips of a spider either zigzagging or hopping across the screen, either towards an obvious goal (i.e., a bug) or to an empty location. Participants were then instructed to describe what they had seen. Gestures that depicted the motion of the spider were analyzed for whether they expressed manner. We hypothesized that the presence of the goal would make manner less salient, so participants would be less likely to express manner in gestures in the goal condition than in the no-óógoal condition. As predicted, participants produced manner gestures on significantly fewer trials in the goal condition than in the no goal condition, $t(14) = 2.48$, $p = .014$, $d = .53$.

However, overall levels of manner gestures were low. Therefore, in Experiment 2, we sought to increase the salience of the spider's movement. We increased the size of the spider's zigzags, and we replaced the hopping manner with spinning. In other respects, the experimental design, procedure and hypotheses were the same as Experiment 1. In this experiment, there was no significant effect of goal on expression of manner in gesture, $p = .46$. Thus, Experiment 2 failed to replicate the indirect effect of goal presence on manner gesture production. However, the zigzag trials in Experiments 1 and 2 offer a direct contrast of perceptual salience. In Experiment 2, the spider's movements were large and perceptually salient, whereas in Experiment 1, the movements were small and subtle. We compared manner gesture production on zigzag trials across both experiments. Participants were more

likely to produce gestures depicting manner in the perceptually more salient zigzag trials in Experiment 2 than in the perceptually less salient zigzag trials in Experiment 1 (84% vs. 35%; chi-square test, $p < .001$). These results are compelling because other factors, including lexical access (i.e., retrieving the word "zigzag") and other task demands (i.e., memory load) did not differ between experiments. Furthermore, perceptual salience affected whether or not manner was expressed in gestures, and not solely the size of those gestures.

In sum, increased perceptual salience of manner led to increased production of manner gestures. These findings support the view that when people observe actions that are highly salient, their representations of those actions become highly activated, and they therefore produce gestures.

Parallel Session 2 – Tuesday, July 8th 2:00 pm – 3:40 pm

Education 1: Talks

The role of gesture in higher order thinking

Rebecca Frausel, Cassie Freeman, Lindsey Richland, Susan Goldin-Meadow

Higher order thinking (HOT) has been identified as essential to children's ability to become academically successful thinkers, and includes making inferences, drawing comparisons, understanding taxonomies, and abstracting away from specific instances. Importantly, HOT is malleable under educational conditions, but the role of early parental input has been largely undocumented. In contrast, parent and child use of gesture before school entry has been implicated in many facets of children's vocabulary learning (Rowe, Raudenbush & Goldin-Meadow, 2012) and may therefore impact the growth of other types of linguistic development, such as HOT language. Moreover, interfering with gesturing leads adults to make fewer inferences (Nathan & Johnson, 2011) and lack of comparative gestures in classrooms is correlated with lower mathematics achievement (Richland et al., 2007). Thus, seeing gesture in adult talk, and using gesture in children's own talk (Ping & Goldin-Meadow, 2010), could reduce the processing demands of HOT, facilitating its development. Little work has explored early gestural support for HOT children receive from their parents, nor gestural support children provide for themselves. This paper describes the use of gesture in parents and children during higher order thinking and other talk across early development.

Participants were 25 parent-child dyads at three time points: 18, 38, and 58 months. Data are from a larger set of 60 dyads, representing the economic, racial, and ethnic diversity of Chicago's monolingual English population in 2000. Family income averaged $56,000 and primary caregiver education averaged slightly less than a Bachelor's degree.

At each point in time, spontaneous interactions between the parent and child were recorded for 90 minutes. Participants were instructed to behave as they normally would; typically engaging in play, meals, etc. Speech was divided into utterances, defined as any sequence of words preceded and followed by a pause, a change in conversational turn, or a change in intonational pattern. Each utterance was coded for whether it was accompanied by gesture, and whether it displayed one of four categories of HOT: Abstraction, Inference, Comparison, and Hierarchy.

While there were significant differences in the number of utterances by parent education level, parent education level did not explain differences in percent of HOT speech in parents or children. However, as children got older, most increased their percent of HOT speech, as well as the percent of HOT speech accompanied by gesture, while adults remained consistent. Multiple regression demonstrated that the percent of parents' HOT speech accompanied by gesture, as well as the total gesture rate of parents and children, significantly affected the percent of children's HOT speech accompanied by gesture; however, parental education did not explain any additional variance.

These results suggest individual differences exist in parents' and children's HOT and gestures' production rates. These differences are unrelated to parent education, and instead depend on the nature of their interactions. As children age, they increasingly utilize gesture during their higher order thinking speech, suggesting gesture may be one mechanism through which children perform more complex thinking. Ultimately, this paper demonstrates the importance of early gesture, both produced and observed, for children's later cognitive development.

Gesture and creativity in children: Untapping creative potential with the hands

Elizabeth Kirk, Carine Lewis

Data will be presented from three experiments that have examined the role of gesture in children's creative thinking. Creativity is most commonly referred to as the ability to generate ideas that are novel, yet useful (Runco & Jaeger, 2012). Physical movement, including hand gestures, can influence idea generation. Encouraging children to gesture has been revealed to encourage them to think about problems differently and explore alternative problem solving strategies (Broaders, Cook, Mitchell & Goldin-Meadow, 2007; Cook, Mitchell & Goldin-Meadow, 2008; Goldin-Meadow, Cook & Mitchell, 2009). Very recently research has explored how the physical experience of gesture impacts upon creative thought processes, however so far this has only focused on adults. For example, Slepian and Ambady (2012) reported that participants who made fluid arm movements (tracing curved lines) scored higher on a measure of creativity (the Alternate Uses Task, AUT; Guildford, 1967), generating significantly more original uses than those who made non-fluid arm movements (angular lines).

We aimed to test the relationship between gesture and creativity in children and explored the extent to which gestures facilitate children's divergent thinking. We expected that children would produce iconic gestures as they generated alternative uses for items (i.e. gestures that convey semantic meaning), and that doing so would benefit their task performance. Iconic gestures provide a means to represent the target items symbolically, and this may help the child generate alternative affordances for that item. Gestures can convey the same basic

idea as speech in a visuospatial rather than a verbal representational format, and this can alleviate some of the speaker's burden (Cognitive Load Hypothesis, Goldin-Meadow, Kelly, Wagner, 2001). Thus, if children gesture when they complete the AUT task, they may reduce their cognitive load and generate a greater number of original uses. Furthermore, gestures may provide a spatio-motoric route to memory and activate associated representations (Morsella & Krauss, 2004). According to the Image Maintenance Hypothesis (De Ruiter, 1998; 2000; Wesp, Hesse, Keutmann, and Wheaton, 2001), gestures can help the speaker maintain mental imagery in working memory during speech production, and this may further serve to reduce cognitive load.

We hypothesized that there would be a positive relation between children's gesture production on the AUT and their creativity scores. This hypothesis was tested in a series of three experiments. In experiment one, we compared the gesture production of children (n = 26, aged 9V11) on a convergent (picture naming) and divergent thinking (AUT) task and found greater gesture production to be associated with enhanced creative thinking. In a second experiment we tested the impact of suppressing gesture on children's creative thinking (n = 50, aged 10-11) and, again, found a positive relationship between gesture production and creativity. A third experiment tested the impact of encouraging children (n = 54, aged 9-11). Children's originality and fluency scores were significantly higher when they were encouraged to gesture whilst performing the AUT. Taken together, our findings clearly demonstrate that gesture facilitates divergent thinking. Gesture thus offers a way to tap into the creative potential of children.

Coordinating talk and object-related gestures in instructional demonstrations

John Rae

The nature of the temporal coordination of a speaker's talk with their hand gestures is a central concern of gesture studies (e.g. McNeill, 1992). However, in some situations, gesturing hands might be occupied with objects that constrain their capacity to gesture - or indeed gesturing might require the acquisition, or movement, of an object, or objects (Goodwin, 2000; Streeck 1996). Moreover, relevant objects are commonly involved in larger courses of action, for example, handling a glass or cup can project taking a drink from it (Streeck, 1995), such that the availability of an object, or the implications of its use, can be a constraint or an opportunity. Consequently, object-related gestures can be subject to specific considerations that arise from the properties of the object in question (see also Brassac, Fixmer, Mondada, & Vinck, 2008).

This present report aims to further examine object-related gestures through analysing their production in a specific setting: instructional demonstrations in craft workshops. Instruction in craft practice makes extensive use of gesture, characteristically with respect to the tools and materials through which the work is done. The sites where craft skills are developed often involve learning through observing or coparticipating with expert practitioners (Marchand, 2008; Sennet, 2008) or through forms of guided participation in activities (Ekström, Lindwall, & Sóljó, 2009).

The present study draws on video recordings of 16 classes in a printmaking studio and a metalworking studio and uses Conversation Analysis to examine the sequential organization of actions within these settings. In particular, I focus on instructional demonstrations, which involve extended displays of how certain tools are used, or how matierals are handled, or how specific artefacts are created.

The analysis firstly shows how instructors accomplish the temporal coordination of talk with object-related manual actions (for example uttering the word "cutting" at the same time at they use scissors to make a cut). Secondly, I present a group of practices that are used by the instructors to achieve this temporal coordination. Two particular practices are contrasted (a) the anticipatory grasping of an object so that it is in-hand and ready for use alongside instructional talk and (b) the delaying of the progress of talk to allow for the acquisition of an object. Finally, a collection of cases where the instructors' talk and gesture is apparently occasioned by something that which occurs in the course of doing something else, thereby modifying one course of action in order to accommodate another one.

The paper concludes by discussing the nature of the talk-gesture relations found in this setting in comparison with those found in re-enactments (Sidnell, 1996) and how the structures of practical tasks (Lerner, Zimmerman, & Kidwell, 2011) can be a constraint and a resource in constructing a multimodal demonstration in a setting concerned with the transformation of objects.

Gestures make abstract science phenomena more concrete for learning

Melissa Singer

Background: When adults and children talk about spatial phenomena and events that are difficult to conceptualize, they often produce hand gestures that bear close resemblance to the image they are describing (McNeill, 1992; Roth, 2003; Hostetter, Alibali, & Kita, 2006).Moreover, students and teachers produce gesture when they talk about abstract science phenomena in the classroom and in laboratory settings (Crowder, 1996; Roth & Lawless, 2002; Singer, Radinsky, & Goldman, 2008).The purpose of this study is to understand the role that hand gestures play in learning abstract concepts in science, in particular, plate tectonics (i.e., a theory of how earthquakes and volcanoes are formed).In my previous work I found that children often produce the correct scientific models in gesture before they articulate them in speech during peer group discussions on plate tectonics (Singer, Radinsky, & Goldman, 2008).However, it was not clear from this re-

search whether or not children changed their scientific understanding or models of plate movements as a function of observing another 's gestural representation or the act of producing one's own gestural representation, or both.In the present study, we extend these findings in several ways: (1) experimental manipulation of gesture input (versus observing naturally occurring gesture input) and (2) assessing whether gesture plays a role in adult learning of science concepts (versus children's learning).

Method: College students participated individually in a short, one-on-one instructional session on plate tectonics.Students were randomly assigned to instruction with gesture (i.e., representational gestures illustrating plate movements and tracing gestures on maps of plate boundaries) or instruction without gesture (i.e., presenting diagrams and maps on a screen without gesture).Students in both conditions were presented with diagrams of plate movements and maps of plate boundaries on a computer screen, the only difference was whether or not they received gesture in the instruction.Students were individually administered five, open-ended questions on plate tectonics both before and after instruction in order to assess their learning.The entire session was videotaped and the students' responses to the open-ended questions were transcribed and coded for plate tectonic concepts/models using a previously developed coding system for plate tectonics (Singer, Radinksy, & Goldman, 2008).

Results: Preliminary results revealed that overall students who were instructed with gesture made correct and specific changes to their models of plate movements in both speech and gesture after instruction compared to students who were instructed without gesture.For example, before instruction some students produced vague hand motions to represent plate movements in response to the open-ended questions (e.g., using one hand to make a circular motion).However, after instruction with gesture these same students refined their models by producing a more specific and correct plate movement in gesture (e.g., indicating with two hands how plates move and converge/diverge in space and time).These results indicate that observing gesture plays a causal role in making abstract, spatial phenomena in science more concrete

Towards a formal description of gesture and the speech-gesture interface: Panel

Speech-gesture-interface constructions for gestures accompanying German verb phrases

Florian Hahn, Insa Lawler, Hannes Rieser

We currently study speech-gesture occurrences where gestures accompany three different classes of German verbs: verbs of motion, verbs of perception and stative verbs. In our talk, we give an overview on our findings regarding the three verb classes and our work concerning how their meanings can be interfaced in order to yield a multi-modal proposition. The empirical basis for our work is a systematically annotated corpus: the Bielefeld Speech-and-Gesture-Alignment-corpus (Luecking, 2013). It consists of 25 fine-grainedly annotated dialogues of dyads engaged in a route-description task about a "bus ride" through a Virtual Reality town.

Regarding the three different verb classes, we have observed the following: While gestures accompanying verbs of motion often depict (the direction of) the path to take, verbs of perception are not only accompanied by gestures representing the object seen or the line of gaze, but also by discourse gestures used to support dialogue structure; as for stative verbs, the accompanying gestures are used to depict the objects described, their location or some of their properties.

We also have found that speech-gesture-overlap can be sensitive to the use of "sentence bracket" ("Satzklammer") constructions. In such constructions the prefix of a verb can be separated from the finite verb stem to be put to the end of the sentence. Prefix and stem together embrace the German "Mittelfeld". One example is the verb "hindurchgehen" ("walk through") in the utterance "Du gehst zwischen den beiden Kirchen hindurch" ("You walk in between the two churches through"). In examples from our corpus the stroke of accompanying gestures extends from the finite verb to the prefix, even when PP- or NP-constructions are integrated between the stem and the prefix.

For analyzing the contribution of gesture meaning to verb phrase meaning, we have developed the following methodology (Rópke et al., 2013): Concentrating on the static semantics of speech-gesture occurrences and assuming that gesture and speech are semantically related, we aim at constructing a multi-modal proposition. We provide first a compositional semantics for the speech part and a compositional semantics for the gesture part. The meaning for the gesture is reconstructed from the SaGA-annotations and the use of motion capturing technology in order to classify the gesticulated shape. For interfacing, both representations are extended adding a parameter in order to compositionally combine them. Subsequently, the extended presentations are fused into the interface proper. Here, the speech representation overrides gesture representation due to scopal considerations. In the end, the interface provides a multi-modal meaning for the speech-gesture occurrence, hence, the idea of a "unified semantics" is maintained. However, due to the workings of the interface procedure, we also get independent semantics for the speech part, the gesture part and the function of the interface. Compositionality is modelled using typed lambda-calculus and ideas from Combinatory Logics. Formal solutions are based on the works of Reichenbach (1947), Montague (cf. Thomason, 1974) and Parsons (1990).

Pointing and iconic gestures in multidimensional semantics

Cornelisa Ebert

We relate the study of multimodal phenomena to discussions about "multidimensional meanings' in formal semantics and pragmatics (see e.g. Potts 2005, 2012). We are interested in the question how information parts from different modalities, i.e. from the gesture and the speech channel, combine and interact, the working hypothesis being that they function very much like pieces of information of different dimensions within speech.

In particular, this talk has the following aims:

1. Relate the interaction of gesture and speech to "multidimensional phenomena" in formal semantics and show that speech and gesture interact in exactly the same way as different dimensions of speech among each other (i.e. "at issue" vs. "non-at-issue" material, Potts 2005).

2. Discuss the role of demonstratives like German "so" ("such a"/"like") and "dies"- ("this") and propose that they act as "dimension shifters' (cf. Bühler 1934 and Diessel 2006, where demonstratives are argued to serve the purpose of calling the addressee's attention and thus create a point of "joint attention").

3. Present a first attempt towards a formal semantic analysis of the interaction of speech-accompanying pointing and iconic gestures with speech.

The prime example showing the existence of meaning components of different semantic and pragmatic dimensions are conventional implicatures, e.g. contributed by appositives as in Paul, a famous horse riding instructor, has called yesterday, or so-called "expressives" such as the adjective damn or nouns like cur. These expressions and constructions are argued to bring in information that is not "at issue" at the time of utterance, but sneaked in as "secondary" information. The noun "cur", for example, denotes the set of dogs, just like its non-expressive counterpart "dog", but brings in a second component expressing that the speaker does not like dogs in general or a specific dog under discussion. This second expressive meaning component is not asserted, but conveyed on a different dimension. And also appositive material is argued to contribute meaning components of this same second dimension. Empirically, this dimension division is reflected among other things by the fact that secondary meanings cannot be the target of negation and cannot be directly denied in discourse. In the above example, the hearer cannot question the part that Paul is a famous horse riding instructor by saying "no, that is not true" or "I don't believe this", but only the proposition that Paul called yesterday.

We will show that speech-accompanying gestures generally behave like such items in most (or potentially all) respects. For instance, they, too, cannot be targeted by negation or denied directly. We will further discuss one notable exception, namely the presence of a demonstrative expression. If there is a demonstrative expression involved, gesture material can be made at issue and thus contribute to the main assertion (cf. Fricke 2012, who shows that so is a means to integrate gesture semantics into speech).

The striking parallels of multidimensional and multimodal phenomena suggest adopting the formal models known from the treatment of multidimensional meanings also to multimodal phenomena. In our talk, we will suggest a first approach towards such a model.

Bridging the gap: Syntactic complexity and recursion between empirical observation, algorithm, and Language Theory

Ellen Fricke

Syntax is still an understudied area in gesture studies and research on linguistic multimodality. Recent studies on this topic that take into account the media-specific properties of articulators (for multimodal integration in noun phrases, see Fricke 2008, 2012, 2013; Ladewig 2011; for syntactic complexity in gestural stroke sequences, see Fricke 2008, 2012; Bressem 2012; for representations of co-speech gestures in unification-based grammars, see, for example, Lücking 2013 indicate that further research is needed in order to gain a deeper understanding of the syntactic structures that characterize each modality and how these may be related across modalities.In the current debates about recursion and linguistic complexity, various positions have been adopted (see, for example, van der Hulst 2010; Zwart 2011; Sauerland and Trotzke 2011). Hauser, Chomsky, and Fitch (2002) assume that recursion is specific to the human faculty of language and is not to be found either in animals or in human cognitive abilities other than the faculty of language. Author (2007, 2008, 2012, in press) has shown that constituency and recursion can be manifested by co-speech gestures alone. Gestural constituent trees and phrase-structure rules based on the analysis of empirical examples reveal the structural property of self-embedding, in that gestural constituents can contain other gestural constituents of the same type. Syntax is still an understudied area in gesture studies and research on linguistic multimodality. Recent studies on this topic that take into account the media-specific properties of articulators (for multimodal integration in noun phrases, see Fricke 2008, 2012, 2013; Ladewig 2011; for syntactic complexity in gestural stroke sequences, see Fricke 2008, 2012; Bressem 2012; for representations of co-speech gestures in unification-based grammars, see, for example, Lücking 2013) indicate that further research is needed in order to gain a deeper understanding of the syntactic structures that characterize each modality and how these may be related across modalities.

In the current debates about recursion and linguistic complexity, various positions have been adopted (see, for example, van der Hulst 2010; Zwart 2011; Sauerland and Trotzke 2011). Hauser, Chomsky, and Fitch (2002) assume that recursion is specific to the human faculty of language and is not to be found either in animals or in human cognitive abilities other than the faculty of language. Author (2007, 2008, 2012, in press) has shown that constituency and recursion can be manifested by co-speech gestures alone. Gestural constituent trees and phrase-structure rules based on the analysis

of empirical examples reveal the structural property of self-embedding, in that gestural constituents can contain other gestural constituents of the same type. In the current debates about recursion and linguistic complexity, various positions have been adopted (see, for example, van der Hulst 2010; Zwart 2011; Sauerland and Trotzke 2011). Hauser, Chomsky, and Fitch (2002) assume that recursion is specific to the human faculty of language and is not to be found either in animals or in human cognitive abilities other than the faculty of language. Fricke (2007, 2008, 2012, in press) has shown that constituency and recursion can be manifested by co-speech gestures alone. Gestural constituent trees and phrase-structure rules based on the analysis of empirical examples reveal the structural property of self-embedding, in that gestural constituents can contain other gestural constituents of the same type.

Within the framework of generative grammar, and admitting Hauser, Chomsky, and Fitch's (2002) hypothesis that recursion is the only defining criterion for the human faculty of language in the narrow sense (FLN), finding recursion in co-speech gestures has the language-theoretic implication that natural spoken languages have to be conceived of as inherently multimodal. Conversely, rejecting the claim that language is therefore fundamentally multimodal implies that recursivity cannot be taken to be the defining criterion of the language faculty in the narrow sense, as the Chomskyan model proposes. Within the framework of generative grammar, and admitting Hauser, Chomsky, and Fitch's (2002) hypothesis that recursion is the only defining criterion for the human faculty of language in the narrow sense (FLN), finding recursion in co-speech gestures has the language-theoretic implication that natural spoken languages have to be conceived of as inherently multimodal. Conversely, rejecting the claim that language is therefore fundamentally multimodal implies that recursivity cannot be taken to be the defining criterion of the language faculty in the narrow sense, as the Chomskyan model proposes.

In view of these findings on recursion and syntactic complexity in co-speech gesture, the following questions will be discussed: What kind of mutual relationships between empirical observation, formal description, and language-theoretical implications can be observed by reconstructing the process of scientific elaboration underlying each approach to inquiry. What kind of interdependencies occur at which step of analysis. Do empirical observations and non-formal descriptions only generate scientific heuristics, as some formal linguists assume. How far can both formal descriptions and observation-based descriptions contribute to heuristics in gesture studies'.What are the specific goals and achievements of each approach. To what extent can formal approaches be used to describe information residing in gesture and multimodal communication'.How can both formal and descriptive approaches benefit from each other. In view of these findings on recursion and syntactic complexity in co-speech gesture, the following questions will be discussed: What kind of mutual relationships between empirical observation, formal description, and language-theoretical implications can be observed by reconstructing the process of scientific elaboration underlying each approach to inquiry. What kind of interdependencies occur at which step of analysis'.Do empirical observations and non-formal descriptions only generate scientific heuristics, as some formal linguists assume'.How far can both formal descriptions and observation-based descriptions contribute to heuristics in gesture studies'.What are the specific goals and achievements of each approach'.To what extent can formal approaches be used to describe information residing in gesture and multimodal communication'.How can both formal and descriptive approaches benefit from each other.

The display situation

Andy Lücking

The talk makes two main contributions: Firstly, display situations are introduced as a situation semantics notion for couching co-verbal gestures. Secondly, two grammatical kinds of speech-gesture combination are distinguished. The co-occurrence of speech and gesture takes place in two kinds of constructions: On the one hand, a gesture can occur contingently, that is, without being required or introduced by some linguistic element. On the other hand a gesture occurrence may be forced. The latter is the case when the gesture has to occur as an object of multimodal subcategorization. Grammar frameworks have to keep these two kinds of affiliation apart, since they correspond to different ways of speech-gesture integration. Contingent gestures do not (or at most to a negligible degree) contribute to the meaning of dialog. Contingent affiliation, however, is the place to capture multimodal well-formedness. Even when they are not a substantial element in the comprehension of the utterance, a mismatch of gesture and speech is usually recognized at least on an early stages of processing (Kelly, Kravitz, and Hopkins 2004). In case of a subcategorized gesture, however, the gesture indeed carries crucial semantic information. As Fricke (2012, Sec. 5.4) argues, a gesture can provide verbally missing content when produced in the scope of a demonstratively used so (German adverb that corresponds to English so, thus, or like this). For instance, an utterance of "The tomato was THIS large" would be semantically incomplete until a value for the dimension of size is indicated by pointing to a reference object or by a "sizing" gesture (Umbach and Ebert 2009). A gesture that provides the information focused by the expression this large is fully integrated into the utterance. In the case of multimodal subcategorization, the gesture event acts like a resource situation (Barwise and Perry 1983). In order to reflect the basic difference that gesture events, unlike resource situations, are situatively produced by the speaker and not exploited in the contingent case, a new kind of situation is introduced: the display situation. A display situation, however, is exploited when subcategorized by a linguistic demon-

strative. This demonstrative typically is a word like this or so, combined with stressed intonation. When put on the "semiotic stage" by such indexical highlighting, the gestural information produced in the display situation is semantically interpreted and gets transferred onto the described situation (see, for instance, the spatio-temporal mappings introduced by Lascarides and Stone (2009)). By means of few examples it is demonstrated how this interaction of type-related, verbal meaning and token-related perceptual gesture classification can be captured within the formal framework of Type Theory with Records (Cooper 2012; Larsson 2013).

Cross-Cultural 1: Talks

Gestural and styles among South African urban black male youth

Heather Brookes

Among urban black South Africans in the Johannesburg region, gesturing is a prominent feature in every day interactions. Gestures frequently depict spoken content, and there is a large vocabulary of quotable gestural forms. Gesturing is most highly elaborated in conjunction with an urban informal slang used among male youth in the teens and twenties when they gather together on the street corners in their local neighbourhoods. Based on ethnographic work involving observation, video-recordings of spontaneous interactions, narrative elicitations and interviews, we compare gestural styles among young men in one community, identify how they vary and analyse their social meanings. Among male township youth, spoken and gestural styles are a vital aspect of gaining status among peers and demonstrating an urban, streetwise and city slick identity. Gestural style indexes local identities and divisions among male social networks. Young men talk about three different male youth identities in the township that are marked by differences in language use and styles of gesturing: softies (those who dont hang out much on the township streets) and their associated subcultures including rappers and bhujwas bourgeois, streetwise township males often referred to as maauthi guys or magents gentlemen and pantsulas ruffians including tsotsis thugs who are engaged in antisocial activities and sometimes crime. These three styles vary in terms of kinesic action, they types of gestures employed, how gestures are deployed in relation to speech, their semantic relation to speech and the use of specific gestural forms. There is also a metadiscourse about gestural behaviour in this community. Gestural behaviour indexes two key social divisions within the township, an urban versus a rural identity and respectability versus disrespectability. These findings on the role of gesture and gestural styles in relation to identity will be discussed in relation to the communicative economy of the urban African environment. The question of what aspects of gesture use are influenced by sociocultural processes will be addressed. These data will also be considered in terms of our understanding of language as a multimodal variable semiotic system.

Máori gestures of the hands, head and eyebrows

James Gruber, Jen Hay, Jeanette King, Lucy Johnston

This paper examines the speech-accompanying gesture and non-verbal behaviour employed by bilingual English-Máori speakers of Máori ethnicity and monolingual English speakers in New Zealand. Physical expression has long been regarded as a key component of Máori oratory and performing arts (Dewes 1975, Matthews 2004, Rewi 2010), and distinctively Máori i non-verbal behaviour in everyday personal interactions has been described by multiple authors (Metge & Kinloch 1978, King 1999, Metge 2005). Our research represents the first quantitative assessment of these observations and intuitions. This paper identifies a set of gestures or actions characteristic of ethnic Máori speakers relative to a comparison group of Paókehaá (New Zealanders of European ancestry). We further consider the effects on the function and frequency of these gestures as influenced by different interviewer ethnicities (Máori or Paókehaá) and by the language being spoken (Te Reo Máori or English).Six bilingual Máori and six monolingual Paókehaá participants completed two sociolinguistic interviews in English, once with a Máori and once with a Paókehaá interviewer. Additionally, the bilinguals performed a third interview in the Máori language with the same Máori interviewer. All participants and interviewers were males under 35 years of age. Interviews were recorded by three cameras; one trained on the participant's head, another on the body, and a third capturing both participant and interviewer. Annotators coded select portions of the full interview for an array of manual and non-manual movement (head, gaze, eyebrows).Our analysis demonstrated at least three behaviours were strongly associated with the Máori participants and performed rarely or not all by Paókehaá: (1) frequent and dramatic use of the eyebrow, (2) a proclivity for use of the head to perform illustrative gestures, and (3) a flat- palm finger-extended handshape to depict certain types of movement path. The interviewer ethnicity and language spoken did not influence the rate at which head gestures or flat-hand depictions of path were produced by the Máori speakers. Raising of the eyebrow, on the other hand, was (a) more likely to occur in the Máori-Máori interactions than with the Paókehaá interviewer and (b) even more likely to occur in the Máori-Máori interactions conducted in Te Reo Máori. Possible explanations for the patterns seen as a function of speaker ethnicity (indexing of ethno-cultural information), interviewer ethnicity (potentially more aligned feedback), and language spoken (the influence of rhythmic and prosodic properties) are discussed and evaluated.Findings from the study offer insight into culturally-grounded differences in gesture and non-verbal behavior within New Zealand, as well as the influence of New Zealand's indigenous language on gestural produc-

tion. More generally, the results contribute to the literature concerning cross-cultural variation in gesture while simultaneously isolating for and examining linguistically-driven differences within individual (bilingual) speakers.

Handedness in Citonga gestures (with cross-cultural comparison)

Karen Sanders, Olanike Orie

Gesture is aptly described as a "backdoor" to cognition (Sweetser 2007). Co-speech gesture has been shown to encode metaphorical source domains (Cienke 1998), aid in the representation of abstract concepts (Perril and Sweetser 2004), and specific handshapes, movements, and directionality systematically structure metaphorical vocabulary in American Sign Language (Taub 2001). Although gesture is a rich source of data for the examination of conceptual metaphor, it is noticeably absent from the critical and political discourse analysis paradigms. In this presentation, I use gesture to investigate which source domains are structuring American understandings of transgenderism, the concept ascribed to those who have begun or completed a change in their sex characteristics from male to female or female to male. Through the examination of twenty transition narratives documented on video, I will show how both co-speech gesture and an emerging lexicon of ASL signs align with spoken and written narrative to support a spatially based representation of gender identity and transition. Recently, there has been a large amount of work analyzing the construction of transgender identity (e.g. Armitage 2008; Valentine 2007), some of which includes linguistic analyses of transgender, transsexual, and drag queen communicative patters (Barrett 1998, 1999). However, there exists no comprehensive analysis of the cognitive models used to understand transgender identity or the transition process. I offer a roadmap for those interested in incorporating evidence from gesture into the identification of unconscious assumptions, which organize speakers' comprehension of complex political topics.

The assignment of gender is talked about and thought about as being located in a bounded region; English speakers qualify and quantify gender and transition through their understanding of movement through space: cross-dressing, transitioning, changing, male-to-female, coming out, intersex. This language is indicative of a binary category model of gender assignment, in which each category is understood as a bounded region in space and transition is a journey with intermediate and final destinations along a path as in (1):

(1) I have often likened my transition to slowly wading out into a cold lake. I take a step or two, shiver a bit at the coldness, and hang out for a bit as my body acclimates. Then I decide if I want to go deeper. All along the transition I have been open to the concept that I can stay where I am, go back, or push deeper. And though several times I have pulled back too deep, too fast, too cold- I have always found myself moving toward transition.

Co-speech gesturing from my corpus, such as two upward facing palms in alternate motion, canonical of decision-making (MAKING DECISION IS WEIGHING), suggests the coming out process is understood as a choice with two alternatives. In one specific example of this gesture, the two palms are subsequently coopted into deictic reference points on the left to right timeline. The temporal reference set up in the gesture signals a spatial threshold, which once passed, cannot be re-traveled.

Semantic change in the visual-spatial modality: Evidence from cross-linguistic body-part namin4

Elise Stickles

In his 1996 response to claims that it is not possible to develop general laws of semantic change, Wilkins investigates changes in body-part naming over a broad cross-linguistic field, and finds that 70% of recorded changes patterned into natural tendencies. However, in order for a tendency to be "natural" cross-linguistically, it must be observed in a typologically varied cross-section of languages. By failing to consider signed languages, Wilkins doesn't take into account an important linguistic family, especially with respect to semantics. Unlike spoken languages, signed languages have highly iconic lexicons (Wilcox & Wilcox 1995). Frishberg's classic (1975) study of diachronic change in American Sign Language (ASL) demonstrates that change in ASL is systematically progressive from iconic representations to more arbitrary forms. Wilcox & Wilcox argue that Frishberg's proposed tendency is limited to the lexical level, and thus propose their iconicity principle, wherein reduction in gestural substance associated with changes from more concrete to more abstract semantics is driven by the physical grounding of the polysemous senses.

This study extends these diachronic analyses of signed languages to Wilkins' proposed tendencies by undertaking a diachronic comparison of modern ASL body-part terms and their French Sign Language (LSF) counterparts with data from their ancestor, Old French Sign Language (OFSL) (Renard and Delaporte 2002). Changes from OFSL to the modern body-part terminology in both languages conform to Wilkins' tendencies, but also follow those of Frishberg and Wilcox & Wilcox. Differences between ASL and LSF primarily entail strategies to accommodating the above-waist signing constraint, as well as culturally-driven differences in genitalia terms. These variations in taboo naming accord with Wilkins' observation that 30% of variation is culturally-specific.

Results of the current study show that signs for external and readily available body parts (e.g., ARM) were more likely to maintain a high level of iconicity via direct deixis, which follows from the iconicity principle. Signs for more abstracted (e.g., BODY) or less visible (e.g., FOOT, HEART) referents are more likely to change into forms that are less iconic, such as fingerspelling or classifiers. These results also accord with

Pyer's (2006) discussion of synchronic ASL body-part terminology, which shows that signs are differentiated via distinctions such as internal vs. external. For example, her observed categorical variations in phonological markedness (Open-B vs. Bent-B handshapes) between surface and internal structures is predicted by the fact that surface body parts are visible and therefore more physically grounded; thus internal-part naming underwent a greater degree of change and acquired a more marked variant. These findings provide a more nuanced perspective on the grounding of semantic changes in physical attributes as well as further empirical support for the original analyses of Frishberg and Wilcox & Wilcox.

By demonstrating that changes in body-part naming in signed languages both conform to patterns predicted by studies of spoken languages and contribute predictable tendencies predicated on modality-specific constraints, this study stresses the importance of signed languages in typological surveys. It also more fully explicates the ramifications of iconicity in the visual modality with respect to semantic change.

Interaction 2: Talks

Gestures in interpreter-mediated clinical interactions: The function of body-directed gestures

Jennifer Gerwing, Shuangyu Li

When physicians and patients do not share a common language, an interpreter can mediate the conversation, assisting with the process of establishing mutual understanding. This process can involve contributions of both speech and gesture. This project aimed to (1) analyze how physicians and patients combine gesture and speech when referring to regions of the body and (2) explore how interpreters reformulate patients' and physicians' gestural or verbal components. Reformulations could distort the meaning of the original utterances, thus having a deleterious effect on the mutual understanding physicians and patients are establishing. If the interpreter preserves gestural components of the utterances (which are observable to the other parties) while transforming verbal components, these distortions could be masked. Studies of gesture use in interpreted clinical interactions are new to the gesture field, and the questions explored here provide a unique means of exploring gesture-speech integration.

We used seven videotaped primary care consultations collected in clinics in the UK that involved either ad hoc or professional interpreters. Physicians were native English speakers, and patients spoke Slovak, Mirpuri Punjabi, and Urdu. All seven consultations were transcribed using Conversation Analysis conventions. Non-English speech was translated into English. For analysis, we used the videotapes and transcripts. Body-directed gestures were those in which the relationship between the speaker's hand and body was intrinsic to the gesture's meaning. First, we located and analyzed all body-directed gestures, which included identifying each gesture's referent and its semantic relationship with the accompanying speech. Second, we analyzed how the interpreter preserved, deleted, or transformed gestural and verbal components of utterances that included body-directed gestures.

The analysis showed that all participants used body-directed gestures with their speech to aid the process of establishing mutual understanding: total = 104 body-directed gestures, distributed as follows: physicians, M=15.1; patients, M=9.4; interpreters M=16.1. Most broadly, body-directed gestures contributed the subject of the utterance (i.e., the topic), while speech contributed the predicate (i.e., something about that topic). Gestures depicted a wide range of body regions and functioned to convey information, demonstrate understanding, and disambiguate interpreted content. Gesture-speech combinations depicted symptom locations (e.g., pain, discomfort), pathways (e.g., stomach acid through the esophagus), and actions (e.g., drinking acid-reducing medication). Interpreters often preserved the gesture, but when they transformed information contributed by speech, they changed the meaning of the utterance, even though the topic was visibly maintained. For example, one interpreter copied the patient's gestures towards his chest area, but transformed his verbal contribution from "it hurts me" and "long term problem" to "all the time pain".

The results suggest that physicians could benefit from attending to patient's and interpreter's gestures in order to check for concordance between the two. However, body-directed gestures are tightly integrated with speech: Even if interpreters preserved visible gestural components of interpreted utterances, transformations of the verbal component could alter utterance meaning.

I see how you feel: Speakers' gestures help people to understand their pain

Samantha Rowbotham, Judith Holler, Alison Wearden, Donna Lloyd

Pain is a frequent feature of medical consultations and must be communicated effectively if health care providers are to understand the experience and provide treatment. However, pain is difficult to verbalise and spoken pain descriptions are subject to misinterpretation (Schott, 2004). It is well known that when we speak we also produce co-speech hand gestures, and during pain communication these gestures have been found to depict aspects of pain that are not contained in the accompanying speech, such as location, sensation and cause of pain (Rowbotham et al., 2012, 2013a, 2013b). Although recipients are known to be able to comprehend the information contained in gestures produced during descriptions of concrete entities and events (see Hostetter, 2011 for a review), it is not yet known whether this is the case for subjective experiences such as pain. We investigated whether untrained observers are able to glean any additional information from the gestures that accompany spoken pain descriptions,

and whether this can be enhanced through a short instruction session on co-speech gestures. Participants (n = 30 per condition) viewed 20 short video clips (mean length = 7.5 seconds) of pain descriptions under one of three presentation conditions: 1) Speech Only, 2) Speech and Gesture, or 3) Speech and Gesture plus Instruction (a short presentation, prior to the video clips, explaining what co-speech gestures are and the types of pain information they can depict). Following each clip, participants provided a free-text description of the pain and a "traceable additions" analysis (Kelly et al., 2002) was used to assess whether participants' descriptions contained any information that was uniquely contained in gestures in the original clips. Participants who had received instruction in co-speech gestures (Speech and Gesture plus Instruction condition) obtained the most information from gestures, while those who did not have access to gestures (Speech Only condition) obtained the least. There were no differences in the amount of information obtained from speech across the conditions, suggesting that neither having access to gestures nor being instructed to attend to these has any detrimental effect on pain understanding. These results suggest that attending to the speaker's gestures during pain communication can enhance the recipients understanding of this subjective experience. These findings have important implications for communication in medical settings, suggesting that health care professionals may benefit from training in co-speech gestures in order to improve their understanding of patients' pain experiences.

Semiotic lexicon: An account of the interactions between bodily semiosis and "episteme"

Rolla Das, Rajesh Kasturirangan

Epistemic negotiations are moment by moment evaluations of the speaker's epistemic states vis-vis her addressee's epistemic states (Stivers et al., 2011; Heritage, 2012). Each turn at talk reveals such negotiations and contributes to the progression of the conversation. In face to face or telephonic talk, these epistemic negotiations are done by various kinds of turn constructional units and are often multisemiotic.

We studied discourse markers, as one of the turn constructional units, which assist in epistemic negotiations. Discourse markers shape the interactional organization of talk by indicating the participants' current epistemic state and orientation towards ongoing talk (Heritage, 1984, 2005; Local 1996; Golato, 2010; Beach, 2005). The markers are multifunctional and polysemous (Fischer, 2006).

The data was procured from a series of natural conversations by Bengali speakers. Investigations of Bengali discourse markers are virtually absent in the literature on multisemiosis and conversation analysis. Our observations of more than 50 hours of casual natural conversation reveals that there are systematic patterns of the multisemiotic ensembles associated with discourse marker use. Specific semiotic forms such as particular head gestures and specific prosodic contours almost always accompany a particular sense of a discourse marker. This projects the possibility of a semiotic lexicon, which speakers might have access to during conversations. While there is a plethora of literature revealing participants dependence on contingent constructions of multisemiotic ensembles (Stivers and Sidnell, 2005), very few studies address the implication of systematicity of ensembles. Systematicity of ensembles indicates the possibility of a semiotic lexicon which has significant implications for our current understanding of interrelations between conceptualisation, embodiment and, communication.

Further analyses reveal that the semiotic lexicon is thematically organized, in other words the ensembles reveal an "onomasiological" structuring (Geeraerts, 2010). It is the similarity of epistemic themes which guarantee a similarity of constellation of the semiotic units rather the lexical units being used as a discourse marker. AcchA and thik Ache, frequently used discourse markers in Bengali, can represent various senses/ epistemic stances. Some of the senses/stances of these discourse markers are significantly distinct from each other while others overlap. For example, AcchA can represent the stance/sense of "acceptance of a proposal". In conversation, this sense is associated with a specific multisemiotic constellation. thik Ache also has a similar sense/stance of representing "acceptance of proposal". During conversation, if participants want to represent "acceptance of proposal" as a preferred stance, they can use either AcchA or thik Ache but the semiotic lexicon used for either of these markers are analogous.

We argue that the semiotic lexicons reveal a rational and cognitive continuity from language being grounded in "bodily experience" to epistemic alignments achieved via distributed multisemiosis. The thematic motivation for the structuring of the semiotic lexicon additionally reveals how semiotically different media converge together on the basis of conceptual semblance and therefore reinforces and extends the cognitive commitment (Lakoff, 1990), a commitment that linguistic organisation reflects general cognitive principles, not specific to language.

Parallel Session 3 – Wednesday, 9th 10:20 am – 12:00 pm

Gesture and spatial reasoning across domains: Panel

Modality affordances in science learning

Cassie Freeman, Kayla Milne, Kensy Cooperrider, Susan Goldin-Meadow

In the science classroom, students are asked to learn from a variety of representational modalities including written texts, pictures, and digital media, among others. One modality that frequently occurs in the science classroom but has been given less attention is the gesture produced along with speech during classroom discourse (Singer, Radinsky & Goldman, 2008). In this paper, we examine whether using different classrooms modalities (diagram, co-speech gesture, typed narrative) leads to differences in students' explanations of a rich spatial concept the process by which the earth has seasons- and, ultimately, their learning of the concept.

Different modalities provide different affordances for a learner's understanding of the scientific concept of seasons. Words can qualify, negate, and hypothesize, diagrams can depict objects and their parts with veridicality and demonstrate fixed relationships between objects (Tversky, Heiser, Lee & Daniel, 2009). Gestures can depict movement, spatial relationships, and changes to those spatial relationships. Diagrams and typed narratives provide a stable record of concepts, while speech and gesture do not. Given these different affordances, we hypothesized that using these modalities would result in differences in how learners'explained the concept of the seasons and, in turn, in how well they understand it.

College students were first asked to explain the process by which the earth has seasons so that another person could learn from it. They were then given a written, multiple-choice pretest about the seasons with items taken from a national survey of science assessments. Next, they read a one-page written description of the seasons, outlining core concepts such as solar radiance, tilt of the earth, and the path the earth takes around the sun. Participants next explained the process by which the earth has seasons so someone could learn from once more, this time in one of four conditions: (1) diagram+speech+gesture (participants were given a pen and paper but not explicitly told to gesture), (2) speech+gesture (participants were told to speak but not explicitly told to gesture), (3) typed narrative, and (4) control (sat quietly for 45 seconds). Finally, participants took a multiple-choice post-test. The explanations participants provided during the first and second explanation phase were coded for the presence of key concepts of the seasons and compared to gains from pre- to posttest.

There are two important results of this study. First, different key concepts were highlighted in the explanations across conditions. Concepts of solar radiance were highlighted in the typed narrative explanations while concepts of the tilt of the earth relative to the sun and the path that the earth takes around the sun were highlighted in diagram+speech+gesture and speech+gesture explanations. Second, there was a main effect of explanation type on posttest gains, with speech+gesture resulting in the most learning. The differences in the kinds of ideas that can be represented by different modalities, as well as the effect of these differences on learning have broad implications for classroom practice.

A double-edged role for gesture in analogical reasoning?

Kensy Cooperrider, Susan Goldin-Meadow

What leads people to use old ideas to solve new problems'. As a window on cognition, gesture may provide insights into the formation of ideas and their transfer from one context to another. As a causal factor in cognition- with demonstrated consequences for learning and memory (Goldin-Meadow & Beilock, 2010)- gesture may also serve as an active ingredient in this formation and transfer. We investigated these possibilities by introducing gesture into Gick & Holyoak's (1983) radiation problem paradigm, a well-known model system for understanding analogical reasoning. First, in the "story phase" of the paradigm, participants (n=78) read two stories, one after another, and described them for a confederate. One story involves a military attack and the other a fire raging out of control, but both feature a problem that is solved by spatially redistributing a force around a central target. Next, in the "similarities phase" of the paradigm, participants described any similarities they noticed between the two stories. In the final phase, under the guise of an unrelated study, participants attempted Duncker's radiation problem, which can be solved by applying the redistribution strategy suggested by the stories. 45 percent (35/78) of participants successfully applied this strategy.

We analyzed participants' spontaneous gestures and speech in both the story and similarities phase to determine whether they predict success on the radiation problem. Gestures in the story phase were predictive of subsequent failure: only 38 percent (15/39) of participants who produced speech-and-gesture descriptions of the redistribution strategy in both stories went on to solve the problem, compared to 89 percent (8/9) of participants who produced speech-only descriptions in both stories and 50 percent (7/14) of participants who produced a speech-and-gesture description in one story and a

speech-only description in the other. One explanation for these results is that redistribution gestures in the story phase- despite being highly schematic in form- reflect or entrench representations that are too grounded in concrete detail to transfer readily to new contexts. Gestures in the similarities phase told a different story: 20/22 (91 percent) of those who mentioned redistribution as a similarity of the stories did so in gesture, and 15/22 (68 percent) of these would go on to solve the problem, compared to 20/56 (36 percent) of those who did not mention redistribution in gesture or speech. When gesturing about redistribution in the similarities phases, participants sometimes recycled or recombined old gestural forms and sometimes created entirely new forms, and we will consider potential consequences of such qualitative differences for analogical transfer.

The findings from this study are correlational, making it impossible to tell whether gesture is itself an active ingredient in analogical reasoning or a correlate of other ingredients. Nonetheless, they are consistent with the possibility that gesture plays a double-edged role in analogical reasoning: by serving to entrench representations, gesture may hinder the formation of the abstractions needed for analogical reasoning; but once an abstraction is in hand, gesture may no longer hinder- and may even help- the transfer of that idea to new contexts.

Gestures reflect and shape knowledge in complex organic chemistry tasks

Raedy Ping, Susan Goldin-Meadow, Mike Stieff

In this line of work, we use gesture to study the role of spatial reasoning in STEM (Science, Technology, Engineering & Mathematics) teaching and learning. We use a pretest-training-posttest paradigm. On each pretest and posttest trial, novices draw stereoisomers of molecules, which requires a basic understanding of chemistry along with spatial transformations of complex molecules. In Study 1, we segmented and systematically coded the information in learners'gestures (and speech) into distinct problem-solving strategies. Some strategies were partially correct (e.g., the entire molecule must be rotated in space: a hand rotates at the wrist in front of the center of the molecule); others were completely incorrect and indicated misconceptions about chemistry (e.g., the length of a line drawn represents the length of a bond: a finger traces the length of a single bond line). Work from developmental psychology shows that children who produce gestures that convey information that differs from the information conveyed in speech (i.e., a gesture-speech mismatch) on a particular task are on the brink of conceptual change on that task (Church & Goldin-Meadow, 1986). We found here that adult chemistry learners who produced gesture-speech mismatches containing partial (as opposed to incorrect) information at pretest were more likely to improve after training than adult learners who produced no mismatches or who only produced mismatches containing incorrect information. Gesture's predictive power could be explained as a signal of increased cognitive instability (Siegler, 2007)a general predictor of change. If so, the content of the gesture should not matter However, we found that only mismatchers whose gestures conveyed partially correct information on the pretest improved after the intervention. Gesture-speech mismatch thus appears to be a marker for cognitive change because it represents implicit information that the learner holds for solving the task at handinformation that can be used by researchers and teachers alike.

In Study 2, we experimentally manipulate the training phase of the study in order to see whether gesture can go beyond reflecting chemistry knowledge to play a role in changing that knowledge. Learners use 3D models to either act out the spatial transformations necessary in solving the stereoisomer task, to gesture about transformations, or to imagine transformations. Gestures may encourage visualization in a way that action does notthey recruit motor resources, but do not provide visual feedback on the outcome, requiring the gesturer to visualize the outcome Gesture is thus more abstract than action, and producing gesture may encourage students to learn at a deeper level. Pilot data from Study 2 suggest that action, gesture, and imagination all help students to learn how to manipulate molecules like those used during the training portion of the study. As predicted, students in the gesture condition show greater improvement from pretest to posttest than do students in the action or imagine conditions on molecules that look different from those used in the training. In other words, gesture promotes the transfer of spatial problem-solving strategies in this organic chemistry task.

Families of gesture: Talks

Pinning down palm-ups

Natasha Abner, Kensy Cooperrider

Palm-up gestures are a commonplace and cross-culturally robust gestural form, attested in diverse speech communities as well as emerging and conventionalized sign language systems, including homesign. However, a clear and predictive account of how palm-up gestures are used has remained elusive, especially in the realm of co-speech gesture. Palm-ups commonly serve as an interactional mechanism that metaphorically "gives", "receives", or "requests" a communicative object (Kendon 2004, Streeck 2009), but they may also be characterized by the pragmatic import of the utterances with which they co-occur. We hypothesized that the core meaning of palm-up gestures lies within the class of "ignorative" (Wierzbicka 1977) utterances: questions, statements of uncertainty, and exclamatives.

To test the association between palm-ups and the above ignorative utterance types, we used a novel gesture elicitation paradigm. Participants (16 dyads) were filmed practicing and performing two scripted dialogues. The dialogues were con-

trolled for the distribution and form of target utterances (six target utterances per dialogue: two questions, two statements of uncertainty, and two exclamatives) but were nevertheless naturalistic in terms of content and the overall flow of the conversation. Participants were told to use their voice, hands, and body to suit their role in the discourse, but were given no instructions about what gestures to produce or when to produce them. This paradigm was successful in eliciting palm-ups and other gesture types from the participants. Using a measure of gestures per 100 words, we found that palm-ups were produced approximately three times as often with target utterances compared to non-target utterances, a pattern which remains uniform across dialogue scripts and target types and is exhibited by all of the participants studied.

An important further question concerns the nature of the association between palm ups and ignorative pragmatics. Do palm-up gestures serve merely to mirror the pragmatic properties of the speech they accompany (the "hand-in-hand hypothesis") or do they play a role in creating the pragmatic import of the utterance (the "trade-off hypothesis") (c.f. De Ruiter, Bangerter, and Dings 2012)'.We found tentative support for the trade-off hypothesis in our first study: target utterances whose pragmatic properties were implicit in the speech (e.g. concealed "the" questions) prompted marginally higher rates of palm-up gestures than target utterances with explicit pragmatic marking (e.g. "wh-" questions). A follow-up study, currently in progress, manipulates this distinction directly by formulating explicit ("What time is it ") and implicit ("The time ") versions of each target utterance. If palm-up gestures fill in to provide pragmatic information where it is only implicit in speech, they may also be able to add it where it is totally absent. If so, we will have an explanation for the pattern in our data: palm-up gestures are strongly, but not exclusively, associated with target utterances and their presence in non-target utterances may endow those utterances with interrogative, uncertain, or exclamative force."

The family of away gestures. Embodied roots of negative assessment, refusal, and negation

Cornelia Müller, Jana Bressem

The paper addresses the emergence of meaning from manual action and the evolution of a gesture family based on a semanticization of a shared effect of the motivating actions of the hand(s). The work presented here, builds upon a range of linguistic, semiotic, and anthropological studies of recurrent forms and functions in co-speech gestures, all of which point out that variations in gesture forms go along with differences in meaning (Calbris 2003, 2011; Harrison 2010; Kendon 2004; Ladewig 2011, Müller 2004; Müller & Speckmann 2002; Payraté.& Teßendorf in press). Kendon's pioneering work on gesture families has shown that such variations may be systematic and constitute gesture families: "When we refer to families of gestures we refer to groupings of gestural expressions that have in common one or more kinesic or formational characteristics. Each family not only shares in a distinct set of kinesic features but each is also distinct in its semantic themes." (Kendon 2004: 227) Departing, in particular, from Kendon's analysis of the Open Hand Prone (OHP) family and further work on "gestures of negation" (Calbris 2003, 2011; Harrison 2010; Kendon 2004), the paper will offer a linguistic and form-based account of a gesture family which is not only based on shared formational features and common semantic themes, but which is additionally motivated by a shared effect of an underlying action-scheme.

The family of Away Gestures was discovered in the context of a data-driven documentation of a repertoire of recurrent gestures of German, based on a corpus of 24 hours of video data from a variety of discourse types. It was identified by applying a linguistic and form based analysis to the motivation of gesture forms (kinesic features and movement gestalts) and their distribution across contexts-of-use.

In presenting results from this research, we will argue that the family of Away Gestures is semantically motivated by the effect of actions of removing or keeping away of annoying or unwanted things from the body. The family has in common that something has been moved away, or something is being kept away from intrusion. Sweeping Away gestures are used to reject and exclude topics of talk, they negate manually. Holding Away gestures refuse and stop unwanted topics of talk. Brushing Away gestures remove and dismiss annoying topics of talk, by rapidly brushing them away from the speaker's body. They assess topics of talk negatively. Throwing Away gestures remove and dismiss topics of talk, by metaphorically throwing them away from the speaker's body. The clearing of the body space goes along with a qualification of the rejected objects as annoying, that is, a topic of talk is being negatively assessed.

It will be concluded that such a form based linguistic approach provides further evidence to a notion of "gestures as visible actions" (Kendon 2004) and offers support for a praxeological understanding of gesture (Streeck 2009, 2013). In short, it opens up a path to systematically reconstruct the embodied roots of gestural meaning.

On the quantification of gestural phrase grouping in lecture-style speech

Stefanie Shattuck-Hufnagel, Ada Ren

The importance of speech-accompanying gesture in the communicative act is increasingly acknowledged (Kendon 1980, 2004; McNeill 1992, 2005; Gullberg 1998, 2008; Mayberry and Jaques 2000; Duncan 2006; and Goldin-Meadow and Alibali 2013, inter alia), as is its connection with the prosodic structure of spoken utterances (Loehr 2004, 2012; Yasinnik et al. 2004; Shattuck-Hufnagel et al. 2007; Neff et al. (2008); Shattuck-Hufnagel and Ren 2010). Less fully explored is the question of how sequences of co-speech gestures are themselves structured into hierarchical groups, as proposed by Kendon (1980). To address this question, we developed a

two-part system for annotating and quantifying manual co-speech gestures. The first set of labels is a perceptual annotation, capturing the viewer's sense of where each stroke ends in time and how sequences of strokes are grouped; this allows computation of the time interval between successive gestures within and between gesture phrasal groupings. The second set of labels is quantitative, based on numerical specifications for six dimensions including hand shape, trajectory shape, and location with respect to the speaker's body. These numerical specifications allow computation of the degree and type of similarity between any two successive gestures.

This annotation system has been applied to a corpus of video-recorded speech from 6 different academic lecturers (6 minutes of speech); annotations are carried out without listening to the speech, since prior work has shown an interaction between gesture and spoken prosody on perception (Krahmer & Swerts 2007). Preliminary results suggest that a) the majority of gestures produced by these speakers in this context are not referential, but instead appear to be examples of what has been called "beats" or "batons" in the literature; b) the majority of these gestures occur in sequences of various sizes that are perceived as grouped together, and c) gestures perceived as belonging to the same phrasal group occur closer in time than gestures separated by a perceived group boundary. However, the perceived grouping may not be due only to the timing; quantitative comparison of dimensions such as hand shape, trajectory shape and location shows that pairs of gestures within a perceived phrasal group are more similar than pairs of gestures crossing a perceived group boundary.

These observations support the hypothesis that gestures occur in groups defined by both timing and similarity in the proposed dimensions, and opens the possibility of comparing the gesture phrase grouping patterns with both 1) the grouping of prosodic constituents, on the one hand (Loehr 2012), and 2) the monologue structure of the lecture, on the other. This approach has the potential to reveal the complex structure of a communicative speech act that includes its morphosyntax (words and sentence structure), its prosody (intonation and timing structure) and its gestural structure. For example, analysis has revealed a striking degree of "clumping" or uneven distribution of these non-referential gestures across the monologue, raising interesting questions about the relationship of these "clumps" of gestures to the speech production planning process.

Practicing gesturecraft: One person's gestures of completion

Jürgen Streeck

The notion gesturecraft (Streeck 2009) indexes a praxeological tradition of studying body motion that originated with Marcel Mauss" essay about techniques du corps but is also informed by Husserl's concept of the "living body" and Merleau-Ponty's "habit-body". Gesture is skilled bodily praxis, comprising symbolic practices by which communicating actors solve a broad and evolving range of sense-making problems. These practices, while not distributed in equal packages among a community's members, are sufficiently shared between individuals to allow gesture to do its work of facilitating shared understanding and coordinating action. In contrast to other approaches, which abstract gesture as a code from the body, the study of gesturecraft (i.e. of gesture as craft) does not treat the maker of gestures as a featureless variable, but as a "concrete subject having habits, interests, and capabilities as a result of accumulated experience" (Thompson 2007: 29). An individual's stock of gesticulation methods is a set of habitualized, personal responses to sense-making tasks that the individual routinely encounters. Each habit-body self-organizes, adopting and adapting pre-made ("cultural", "conventional") while also finding new ("idiosyncratic") solutions to routine communicative demands.

To investigate gestural habitus, we must turn our attention to the individual. This paper presents observations about a family or families of gestural forms enacted by one person in a recurrent type of interactional context: during the completion of a course of action, either manual, communicative, or both. While the gestures share certain features of form (e.g. closing and/or upward rotation of the open hand), their shapes also emerge from physical features of the actions the hand has previously been engaged in. It is shown how the gestures recruit familiar motor-patterns of action (Calbris 2011) and elements of context (the stage the physical activity has reached, the gesture's timing with respect to the unfolding conversational sequence; Schegloff & Sacks 1973) and is coordinated with other body-motions (shifts in gaze; Rossano 2012) to convey a sense of closure.

The habitualization of this gesture family in the communicative idiom of the subjects' an auto-mechanic is interpreted within the context of routine communication tasks that he encounters during his work-days, and as a case in point to illustrate that it is productive to conceive of gestures, not as a shared code, but as a craft that is attentive to the craft and habits of others, yet the result of an individual body's specific self-organization. Data for the study are a video-recording of one work-day in an auto-shop.

Emerging gesture and sign: Talks

Different biases in comprehension and production: Evidence from elicited pantomime

Matthew Hall

All natural languages evolve devices to communicate who did what to whom Elicited pantomime provides one model for studying this process, by providing a window into how humans (hearing non-signers) behave in a natural communicative modality (silent gesture) without established conventions from a grammar To the extent that their behavior is consistent across individuals, dissociable from their native lan-

guage, and reminiscent of patterns in naturally-evolving systems (e.g. homesign, emerging sign languages), hearing non-signers can provide evidence about the cognitive forces that shape emergent language structure.

Previous work in this paradigm has focused on regularities in production, especially constituent order. Pantomime producers tend to use Agent-Patient-Action order (here termed SOV) when describing canonical transitive events (e.g. a woman pushing a box), but they avoid SOV when describing reversible transitive events (e.g. a woman pushing a boy) Among their alternatives to SOV is OSV: BOY WOMAN PUSH.

For reversible events, both SOV and OSV utterances present a potential comprehender with a Person-Person-Action sequence Most accounts assume that this creates ambiguity, leading to communicative difficulty that feeds back to change production behavior However, direct tests of this assumption are lacking; the present study addresses this gap.

When confronted with Person-Person-Action (PPA) gesture strings, comprehenders might have unreliable or inconsistent preferences Alternatively, comprehenders might assume that the first-mentioned person is the agent, or that the agent is determined by the form of the action gesture. Identifying these comprehension heuristics and their relation to production heuristics is an important step in understanding how new human communication systems conventionalize

Experiment 1 presented hearing non-signers with pantomimed strings that varied only in their constituent order (PPA, PPA, & APP) Participants chose which of two pictures best matched the pantomime string The two pictures varied only in their thematic role assignment.

However, naturalistic pantomime generally includes additional cues to thematic role assignment such as spatial indexing, role-shift, and/or proto-verb-agreement. These cues were intentionally omitted from the stimuli in Experiment 1. If comprehenders are sensitive to these cues, they may be more able to discern whether a producer intended a PPA string to mean SOV vs. OSV. Sensitivity to these additional cues might even override constituent order information entirely.

Therefore, Experiment 2 tested PPA and PAP events, but the confederate always produced the two nominal gestures on contrasting sides of space, and we manipulated the form of the action gesture such that it was produced from the perspective of either the 1st-mentioned or 2nd-mentioned nominal. If comprehenders are sensitive to these additional cues, we would expect a different pattern of results from Experiment 1.

Results were highly similar across both experiments: comprehenders in Experiment 1 had an Agent-First heuristic that was as strong for PPA strings as it was for PAP strings. The additional cues in Experiment 2 weakened this bias only slightly.

These results suggest a new perspective on the ambiguity problem in comprehension: it may not be that comprehenders are uncertain, but rather that producers are not acting in accord with comprehenders' general preferences.

A kinematic model for constructed dialog in American Sign Language

John McDonald, Rosalee Wolfe, Jerry Schnepp, Jorge Toro

Avatars that portray sign language hold great promise for improving Deaf / hearing communication in their application in education and automatic translation. Although avatars can portray legible sign language from phonemic tags alone (Wolfe, Cook, McDonald, & Schnepp, 2011), producing believable, natural motion requires the model to also compute a number of sub-phonetic movements that are not included in linguistic annotation. These include subtleties in individual joint timings which, while not part of the linguistic structure, are essential to achieving lifelike, communicative animations of the discourse. These subtleties come in several forms including the timing of anticipatory joint motions that cascade through the body as muscles contract, and compensatory motions that facilitate the maintenance of balance. In fact, the referential shift for constructed dialog is an important case study for such processes, because it affects the entire body.

Referential shifting is a narrative technique often used in American Sign Language (Metzger, 1995; Padden, 1986). To cue constructed dialog, a signer uses a referential shift to assume the role of a protagonist in the discourse (Lillo-Martin, 1995; Lee, Neidle, MacLaughlin, Bahan, & Kegl, 1997; Quer, 2011). The movement depends on the referents' locations in sign space, and their interaction including dialog (Wilbur & Patchke, 1998). For purposes of automatic translation, this is best synthesized via a kinematic model rather than resorting to prerecorded or captured clips. Any synthesis system must support referential shifts in a way that does not interfere with other co-occurring linguistic processes.

The purpose of such a model is to synthesize the loci and timing of constructed dialog from linguistic tags. These include referents, eye gaze, eyelid aperture and other associated nonmanual signals (Rogers, 2011). The model must accommodate all three representational spaces described in (Morgan, 1999), including narrator space, fixed referential framework and shifted referential framework. In particular, the last of these requires the computation of torso movements necessary to orient the signer to indicate both protagonists in the dialog.

The need for these features is widely recognized in the animation industry and professional animators are highly adept at incorporating them in hand animation. Unfortunately, this is time consuming even for experienced animators. Automating the computation of these features would result in more efficient synthesis. Our new model is general enough to support both sign language animation and gesture that accompanies spoken discourse.

Our model's application to ASL is based on studies of annotated corpora (Neidle, 2001; Poor, 2008), and infers the necessary joint orientations from linguistic tags to create the

referential shift movement. It automatically computes staggered timing and compensatory joint motion. The model also cooperates with other co-occurring linguistics processes such as verb agreement, lexical modifiers, and formation of questions. Output from this kinematic model is visualized via a signing avatar. In our presentation we will discuss the details of the model and report the results of our acceptability testing with the Deaf community.

Transitivity and gesture

Suwei Wu, Alan Cienki

Various topics concerning the relation between grammar and gesture have been explored in recent years. For example, Duncan (2002) and McNeill (2003) found out that gesture can reflect verbal aspect; Harrison (2009) and Cienki (2009) argue that there is a variable relation between gesture and modality, like counterfactuals, possibilities, etc. Transitivity, however, which involves various aspects of grammar in a clause (e.g. telicity, punctuality, realis, etc.) (Hopper and Thompson 1980, 2001), has not been explored in this respect. The present study will try to contribute to the discussion on the relation between gesture and grammar with a specific focus on transitivity.

According to Hopper and Thompson, transitivity is a continuum based on ten parameters such as participants, kinesis, telicity, punctuality, realis, individualization of object, and so on. What is more, based on this scale, various utterances have various degrees of transitivity. The transitivity level of "Pat hit Lee" is much higher than that of "Pat likes Lee", though both clauses have a transitive verb, while the transitivity level of "Pat sits behind Lee" is lower than that of the first example but higher than that of the second example, though this clause has an intransitive verb. Our research goal is to explore whether and if so, how different levels of transitivity relate to gesture.

As part of a larger PhD project, this study involves analysis of a set of interviews in American English from the Red Hen video data corpus (https://sites.google.com/site/distributedlittleredhen/home). Firstly, the transitivity of 632 clauses from five interviews in this corpus were coded by hand and classified into different levels on the basis of the transitivity scale parameters: high-transitivity clauses (8-10 high-value paramters), medium-transitivity clauses (5-7 high-value parameters) and low-transitivities clauses (1-4 high-value parameters). Generally, transitive clauses with an action verb have comparatively higher transitivitiy, while intransitive clauses have medium transitivity, and transitive clauses with non-action verbs have very low transitivity.

Secondly, the relation between clauses with various levels of transitivity and gestures like the gestural form (e.g. stroke duration, hold duration) and function (e.g. syntactic and semantic function) is examined. Twenty interviews (each about 5 minutes long) from the Red Hen corpus were coded in ELAN software for gesture forms (phases, handedness, hand shape, orientation, location and movement), the co-ocurring clauses and their transitivity level. Gesture functions (syntactic and semantic function) in the accompanying clauses were coded as well. Preliminary results suggest that there might be a relation between transitivity and gesture. For example, high transitivity clauses (e.g. "Pat hit Lee") tend to have longer stroke duration than low transitivity clauses (e.g. "Pat likes Lee") do; gestures with high transitivity clauses tend to show the action (e.g. "He broke the vase") more often than other clauses do, while those with intransitive clauses (e.g. "He walked along a river") tend to show the motion or path (sometimes also the manner of motion) more often than others do. However, the relation between transitivity and gesture appears to be a complex one relating to both syntactic and semantic factors. For example, in intransitive clauses with embodied actions, like dancing or swimming, the gestures also show the action.

Development 1: Talks

Caregiver's verbal and non-verbal scaffolding effect on children's learning in tasks varied with difficulty

Wing Lam Chong, Shumeng Hou, Ka Ho Choi, Wan Chi Wong

Abundant research has shown that caregivers'speech and gestures can scaffold children's learning. This study explores whether caregivers'instructions would scaffold their children's learning in tasks with different levels of difficulty, and which type of instructions(speech, gesture, or combination of both) is more effective in facilitating children's learning. We asked 3-year-old Chinese children and their caregivers to participate in the present study. They were randomly assigned to two conditions (intervention and control). Each condition had two 3-episode trials. There are two kinds of puzzle tasks, one with 12 pieces and another with 20 pieces. Children played one puzzle in one trial. In the first and third episodes in both conditions, children played the puzzles alone. In the second episode in the intervention condition, children received instructions from their caregivers. Children in the control condition did not receive any instructions. We found that caregivers produced more verbal instructions, but no more instructive gestures in the twenty-piece puzzle game than in the twelve-piece puzzle game. However, the scaffolding effect was only found in the twelve-piece but not in the twenty-piece task. Moreover, only instructions including both speech and gestures (but not those with speech only and gestures only) facilitated children's learning in twelve-piece puzzle task.

Preference for action over perceptual signs by children and caregivers in sign language acquisition

Gerardo Ortega, Inge Zwitserlood, Beyza Sūmer, Asli Özyürek

The structure of iconic signs is motivated by the form of entities and events they refer to. They display an array of strategies to depict a single referent, for instance, by describing the perceptual features of an object or showing an action associated with it (e.g., a toothbrush can be represented by an extended index finger depicting its elongated shape or with a closed fist showing how it is held). Different sign languages exhibit in their lexicon systematic preferences for one of these strategies. Such preferences have been attributed to typological differences2. Here we ask whether, within a sign language, deaf children prefer one variant over the other, and whether adult choice of one lexical variant changes from adult-child to adult-adult interactions.

Five groups of deaf users of Turkish Sign Language (Türk ışaret Dili,TID) took part in a picture description task: pre-school age children (n = 10, mean age: 5;02, range: 3;5 - 6;10), school-age children (n = 10, mean age: 8;03, range: 7;02 - 9;10), parents of the pre-school age children (n = 9, mean age: 30;11, range: 27;3 40;11), parents of the school-age children (n = 9, mean age: 37;11, range: 34;0 46;5) and a different group of adults (n = 10, mean age: 32;8, range: 18;5 45;10). Participants described pictures featuring two inanimate objects configured in a spatial relation (e.g., pen on paper). The two child and the adult groups described pictures to a deaf adult; the two parent groups to their own children. For each spatial description we identified the lexical items labelling the objects. We selected 14 pictures containing objects for which TID has both action and perceptual variants in the lexicon (toothbrush, cup, pen, bathtub and bed). The analysis showed that while adults signing to an adult have a clear preference for perceptual-based signs, children overwhelmingly produced action-based variants. Parents signing to their children produced the same proportion of action and perceptual- based signs. Compared to adults signing to another adult, both groups of parents produced a higher proportion of action-based variants.

The present study suggests that apart from typological differences the different iconic strategies used to depict a referent may be part of the language acquisition process. The match between motoric representations and signs might play an important role during language development because they provide children with a direct link with a concept. The higher proportion of action-based variants during child-adult interactions suggests that parental lexical choice may be a communicative scaffolding strategy to facilitate children's learning of linguistic labels, linking them with their direct experiences with the world. Our results parallel findings from cross-linguistic studies in which hearing children show a marked preference for action over perceptual co-speech gestures when producing a spoken label for a concept. They corroborate increasing evidence that iconicity might play a role in the acquisition of sign languages and that type of iconicity might be a mediating factor in the process.

How do children and adults gesturally manage the activation status of referents in discourse?

Kazuki Sekine, Sotaro Kita

This study investigated how adults and children use speech and gesture to clarify whether a referent is active at a given moment in discourse. As Chafe (1987) argued, a referent that is newly introduced in a story becomes an active referent. Although the referent becomes a semi-active referent as the story moves on to other topics, if the referent is mentioned again, it is reactivated. However, it is difficult to obtain direct evidence that a certain referent is semi-active when analyzing only spoken expression. It is also not clear when the ability to manage a semi-active referent develops. We propose that a semi-active referent is visible in two-handed gestures in which one hand depicts/indicates an active referent, referred to in the concurrent speech, and the other hand that is held in the air indicates a semi-active referent.

The participants were twenty native speakers of English consisting of four different age groups. Each group had five participants, 3-year-olds (M = 3:8), 5-year-olds (M = 5:7), 9-year-olds (M = 9:4) and adults (M = 27.6) (This is a re-analysis of existing data, reported in Özyürek, et al., 2008). A set of 10 video clips depicting motion events was used to elicit speech and gesture. Gestures were coded one of the three following categories; both handed gesture, single handed gesture, or single handed gesture with a semi-active-referent hold (one hand is depicting or indicating the active referent while the other hand is held in the air to indicate a semi-active referent). An ANOVA was conducted on the proportion of the gestures in each category with age group as the between-subject factor. The main effect of the age group was found in the proportion of the single handed gesture with a semi-active-referent hold, $F(3, 19) = 4.11$, $p = .001$, $2 = .50$. Post-hoc tests (Tukey, $p < .05$) revealed that the proportion was significantly higher in adults (20%) than in 3-year-olds (3%) and 5-year-olds (0%), but not in 9-year-olds (5%). However, no main effect of the age group was found in the proportions of the both handed gesture (A=23%, 3yo=33%, 5yo=13%, 9yo=14%) and the single handed gesture (A=57%, 3yo= 65%, 5yo=87%, 9yo=81%).

An ANOVA was conducted on the proportion of the gestures in each category with age group as the between-subject factor. The main effect of the age group was found in the proportion of the single handed gesture with a semi-active-referent hold, $F(3, 19) = 4.11$, $p = .001$, $2 = .50$. Post-hoc tests (Tukey, $p < .05$) revealed that the proportion was significantly higher in adults (20%) than in 3-year-olds (3%) and 5-year-olds (0%), but not in 9-year-olds (5%). However, no main effect of the age group was found in the proportions of the both handed gesture (A=23%, 3yo=33%, 5yo=13%, 9yo=14%) and the single handed gesture (A=57%, 3yo= 65%, 5yo=87%, 9yo=81%).

Novel multimodal behavior of a human-reared gorilla

Nathaniel Clark, Marcus Perlman

While researchers of human gesture have long emphasized the importance of a multimodal approach, comparably little attention has been paid to multimodality in non-human primate communication (Slocomb et al., 2011). Here we document several types of novel multimodal behaviors performed by a human-reared gorilla, and discuss their implications for primate gesture studies and the evolution of human communication. Koko is a 42-year-old female western lowland gorilla (G. gorilla gorilla) who has lived closely with humans since the age of six months, including explicit sign language instruction over much of her life. In the course of her daily activities, she is routinely recorded during interaction with her primary caregivers, which comprises a corpus of several hundred hours. From this video corpus we identified several types of novel (i.e. learned, not species typical) behavior performed by Koko, which involved voluntary control over her breathing and vocal tract, typically in combination with manual gestures and behavioral routines. These behaviors spanned seven different types of behaviors with more than 400 tokens spread across 125 bouts. For example:

- Object-oriented huff: glottal fricative directed at held object, optionally voiced
- Intransitive blow: rounded, voiceless bilabial fricative, accompanied by manual gesture towards her mouth
- Raspberry: voiceless linguolabial fricative with folded tongue protruding through lips
- Mock cough: voiceless glottal plosive with hand at mouth
- Mock blow nose: voiceless nasal fricative achieved by manipulating nostril aperture with hand
- Phone: glottal fricative directed at a phone-like object cradled against ear and cheek
- Clean glasses: open-mouthed voiceless glottal fricative directed at eyeglasses held to mouth

The seven behaviors we catalog have important implications for claims related to the degree of flexibility in great ape communication through the vocal compared to the manual modality (e.g. Tomasello, 2008). They reveal learning and voluntary control in the coordinated actions of a large array of both vocal and manual articulators: hands and fingers, the breathing apparatus, the larynx, and supralaryngeal articulators including the lips, tongue, and velum. Except for the raspberry, which has a relatively complex supralaryngeal articulation, all tokens were accompanied by manual gestures and actions. In these behaviors, Koko makes use of several dimensions of articulatory parameters employed in the phonology of spoken languages, including voicing (voiced and voiceless), place (labial, linguolabial, glottal), manner (stop, fricative), lip roundedness (rounded, unrounded) and nasality (present or absent). Collectively, these behaviors illustrate that Koko has some degree of independent control over each of these articulatory parameters, to a much greater extent than commonly assumed in primate communication research that tends to stress flexibility in manual communication. Researchers are currently developing a multimodal approach to language evolution (e.g. McNeill, 2012), and these data provide a point of comparison to non-human primates. Specifically they suggest that the evolutionary underpinnings of multimodal human communication may be shared among extant great apes, and would therefore have been present among ancestral hominids.

Brain 1: Talks

Gestures make memories, but what kind? Evidence of dissociations in amnesic and Parkinson's patients

Nathaniel Klooster, Susan Wagner Cook, Ergun Uc, MelissaDuff

Extensive work has documented the power of hand gesture to assist in communication, reduce working memory load, and drive new learning. This work has shown that for the listener, seeing speech-accompanying gesture facilitates later recall, and for the speaker, gesturing at either encoding or recall facilitates memory performance. However, the cognitive and neural mechanisms supporting these effects are largely unknown. To investigate the memory system that supports gesture's effects on learning and memory, this study examines gesture production and gesture comprehension in 1) a group of 6 patients with bilateral hippocampal damage and severe declarative memory impairment but intact procedural memory (amnesic patients), 2) a group of 12 mild-moderate, non-demented patients with Parkinson's disease with intact declarative memory, but impaired procedural memory, and 3) a group of 39 healthy comparison subjects individually matched to each patient, with intact declarative and procedural memory. To examine gesture production, subjects first solved the Tower of Hanoi (TOH) either on a physical board with discs, or on a computer with a mouse, before explaining their solution. The healthy and amnesic groups gestures reflected their prior experience; after solving on a physical board, they produced higher, more arching gestures in their explanations than after solving with a mouse. The Parkinson's group failed to show a difference in their gesture production across conditions. To test gesture comprehension, subjects first viewed videos of someone explaining how to solve TOH, and then solved the task themselves on a computer. The videos featured either large, arching hand gestures, or flat, sideways gestures. The mouse trajectories subjects used in their solutions were recorded and analyzed to look

for evidence of which explanation they viewed. Hippocampal amnesic patients behavior was again indistinguishable from healthy comparisons; these subjects used significantly higher, more curved mouse trajectories when solving the task on a computer after seeing the high, arching gesture video than after seeing flatter gestures. The Parkinson's group showed the effect but the magnitude was significantly smaller than the other group's. Taken together, these results suggest that the power of gesture to influence communication and learning can be supported by non-declarative memory substrates, independent of the hippocampal declarative memory system. Thus, gestures may offer a new window through which to study and treat patients with declarative memory impairment, including those due to Alzheimer's Disease (AD), Traumatic Brain Injury (TBI) or the normal healthy aging process. Gesture is a large domain of behavior in which to look for signs of previous experience in memory impaired patients, memories that may no longer be accessible to declarative memory or reportable through speech. Combined with the findings of the powerful role gesture can play in driving new learning, our findings suggest that gesture may be a potential target for rehabilitation or intervention approaches. Finally, because gesture and speech are so tightly integrated during communication in the healthy brain, the study of gesture may offer an exciting window into interactions of declarative and non-declarative memory systems during ongoing communication and cognition.

Gestures as test ground for the social neuroscience of Peirce's Universal Categories

Linn-Marlen Rekittke, Dhana Wolf, Klaus Mathiak, Irene Mittelberg

Charles Sanders Peirce (1960) assumed three Universal Categories (henceforth UCs) to underpin all processes of perception, reasoning, and communication: firstness (e.g. quality, potentiality of meaning), secondness (e.g. facts, contextualized meaning), and thirdness (e.g. habits, patterns, rules). Feeding into Peirce's more specific sign-object relationships (i.e. icon, index, symbol), the UCs interact in each dynamic semiotic process, with one of them determining the predominant local function of a given bodily sign (Enfield 2011; McNeill 2005). Spontaneous co-speech gestures are particularly apt at testing the validity of Peirce's UCs, since they are not as highly codified as thirdness-driven spoken or signed languages; they may thus exhibit all three UCs to greater and more varying degrees. The goal of this study is to utilize Peirce's UCs for gesture research regarding: 1) the annotation and analysis of co-speech gestures data, and 2) the investigation of the neuro- cognitive representation of gesture understanding with functional magnetic resonance imaging (fMRI). This paper presents the theoretical framework and first results of a larger- scale research project at the junction of semiotics and social neuroscience. Here, the focus is on firstness and thirdness and how they may determine the way gestures are interpreted in two conditions, e.g. with and without speech. For data elicitation, 14 German native speakers were asked to retell and reflect on three short movies of different perceptual qualities and contents: M1 focuses on sounds and family conflict; M2 on relationships and taboo; M3 on dance and the meaning of life. Our findings provide first cues on the interaction of the UCs in gesture interpretation. Without speech, observers could recognize and sense certain qualities in the gestures, considering several potential meanings. Some gestures remain firstness-dominant with no clear form-meaning relationship even when considering the speech content; others conveyed meaning without speech due to their thirdness characteristics here seen as entailing, for instance, conventional codified meanings such as conveyed by emblems (Kendon 2004), but also more abstract meanings such as evoked by pragmatic or recurrent gestures (Muller 1998, 2010), underlying image schemata and metaphors (Cienki 2010), or common metonymic references to the speaker's body (AUTHOR, in press). In the second part of the paper, we will discuss how these cognitive-semiotic processes may be reflected in distinct brain activation patterns (Andric 2012). Gestural stimuli embodying Peirce's UCs may differentially involve the three processing streams of visual information (Binkofski 2012): the ventral stream (encoding object properties) for firstness; the dorso- dorsal stream (encoding spatial relations) for secondness; and the dorso-ventral stream (encoding action understanding) for thirdness. Video sequences from phase 1 of this study were presented as stimuli in an fMRI experiment. Here, too, the presentation of co-speech gesture with and without speech allows the investigation of how potential and contextualized meaning is processed. Based on imaging data of 20 healthy subjects (10w) we will present preliminary results. Differential involvement of the dorso-dorsal and ventro- dorsal stream, as expected for secondness and thirdness, will be discussed in light of the proposed dorsal-stream division (Rizzolatti 2003).

Neural responses during perception of naturally produced, meaningful co-speech gestures

Jill Weisberg, Amy Hubbard, Karen Emmore

Meaningful manual gestures are an integral component of spoken language perception1, yet little is known about how the brain integrates these gestures with the co-produced speech signal. Previous studies presenting manipulated gestures made by actors in carefully scripted stimuli revealed that meaningful gestures increased activity in inferior frontal and posterior lateral temporal cortices2. In addition, perception of natural, prosodically-linked co-speech gestures (i.e., beat gestures) implicated the planum temporale as a possible locus of multisensory integration3. However, it is unclear whether similar results occur for meaningful gestures. For example, does the semantic content of meaningful gestures shift the locus of integration from a sensory association region to a more posterior semantic processing region'o Using

fMRI, we examined this question using naturally occurring, unscripted audiovisual communication. We also investigated whether knowledge of a visual-gestural language (i.e., ASL) impacts neural responses to co-speech gesture.

During scanning (GE 3T gradient-echo EPI scans, 40 3.5mm axial slices, 2s TR, 22.4cm FOV), 14 hearing, English-speaking monolinguals (7 female, mean age = 25.14) and 14 hearing native ASL-English bilinguals (7 female, mean age = 22.36) watched unscripted videos of a native English speaker, visible from the hips to the neck. Each of 18 clips (14-19s each) selected from 3-hours of recorded free-flowing speech contained 2-7 meaningful gestures (meaningfulness determined by ratings from an independent group of 20 subjects). Participants passively viewed each clip under one of four conditions: 1) BOTH: co-speech gesture with concurrent speech; 2) GESTURE: co-speech gesture without audio; 3) SPEECH: motionless speaker with speech signal; 4) STILL: the motionless speaker without audio.

Using AFNI, individuals' fMRI response estimates and corresponding t-scores for contrasts of interest (generated by 3dREMLfit) were entered into group-level 3dMEMA analyses.

We found no reliable group differences, suggesting that the neural system supporting co-speech gesture perception is robust and not altered by life-long sign language experience. Across groups, clips of speech with gesture (vs. either alone) elicited increased activation in regions associated with perception of biological motion (bilateral STG), language (IFG), and action (precentral gyri). No regions responded more during speech alone than during speech with gesture. However, regions demonstrating increased activity for gesture alone, relative to speech with gesture, included, most notably, bilateral inferior parietal cortex (BA 40), perhaps indicating less effortful processing of meaningful gesture in the presence of related speech. Critically, in bilateral posterior STG (inferior to the planum temporale), the response to speech with gesture exceeded the combined responses of speech alone and gesture alone, implicating pSTG as an important site for multimodal language integration of meaningful gestures with accompanying speech. Our findings show that perception of natural, meaningful gestures accompanying speech enhances activation in a network of sensory, language, and action processing regions. Further, the locus of multisensory integration for meaningful gestures with speech may be distinct from that of co-speech beat gestures.

Behavioral and neurophysiological correlates of communicative intent in pointing gestures

David Peeters, Mingyuan Chu, Judith Holler, Peter Hagoort, Asli Özyürek

In everyday communication, index-finger pointing gestures are often used to establish triadic joint attention on a referent (e.g., Kendon, 2004; Kita, 2003; Tomasello, 2008). Although such pointing is generally considered a joint action, it is unclear whether and how characteristics of the joint act shape the kinematic properties of the gesture. The present study experimentally manipulated the gesturer's communicative intent as one possible factor influencing the kinematic properties of the pointing gesture, by varying the gestures informativeness. In addition, a first step was undertaken towards understanding the neurophysiological mechanisms underlying the planning and production of referential pointing gestures.

Twenty-four participants pointed for an addressee at one of four circles that lit up on a computer screen, while their three-dimensional hand movement kinematics and electroencephalograms were continuously recorded. The addressee looked at a corresponding screen placed back-to-back with the participant's screen and either saw the same circle light up or did not see a circle light up. This was mutually known by participant and addressee, rendering the participant's pointing gesture either redundant or informative. In the informative condition, participants significantly lowered the velocity of the gesture's stroke movement and significantly prolonged the duration of its post-stroke hold-phase. There was no behavioral difference in the duration of planning the gesture in the two conditions. However, response- locked event-related potentials (ERPs) time-locked to gesture onset showed a significant effect of informativeness in the 100-ms time-window before the onset of the gesture. In addition, stimulus-locked ERPs showed a P300 effect that was significantly more negative for informative compared to non-informative gestures.

In line with findings on instrumental actions like reaching and grasping (Sartori et al., 2009) and the spontaneous production of iconic co-speech gestures (e.g., Gerwing & Bavelas, 2004), the kinematic form of a pointing gesture is influenced by the gesturer's communicative intent. The response-locked ERP results resembled the readiness potential with a slightly more anterior distribution. This suggests an interaction between planning the execution of a motor program and the activation of the theory-of-mind network. Thus, the results of our study fit well with models of speech and gesture production that incorporate communicative intentions, such as the Sketch model (De Ruiter, 2000) and the Interface model (Kita and Özyürek, 2003). In addition, the P300 findings suggest that participants applied greater amounts of attentional resources in planning a more informative gesture (e.g., Polich, 2007). Follow-up studies are carried out to shed more light on the influence of speech-gesture interaction on the form of deictic gesture and speech.

Parallel Session 4 – Wednesday, 9th 2:00 pm – 3:40 pm

Language as multimodal phenomenon: Panel

The bridge of iconicity: From a world of experience to the experience of language

Pamela Perniss, Gabriella Vigliocco

When we consider language in the context of face-to-face communication, two important observations are that communicative expression comprises the use of different channels in systematic and orchestrated ways (e.g. speech and gesture in spoken languages, McNeill 1992; manual signs and mouth actions in sign languages, Sandler 2009) and that language offers widespread opportunity for iconicity, a resemblance between form and meaning. Thus, co-speech gestures may be iconic of properties of their referents (e.g. repeatedly tracing a circle in the air to represent the spinning motion of an entity), prosodic modulations of speech may provide iconic cues to meaning (e.g. the vowel lengthening in loooong to mean very long), and on a lexical level iconicity is found to a high degree across sign languages, but also in many spoken languages (often referred to as sound symbolism), especially outside the Indo-European family.

Nonetheless, iconicity has traditionally been considered to be a marginal and irrelevant phenomenon for our understanding of language (Newmeyer 1992). Rather, the arbitrary and symbolic nature of language has long been taken to be a design feature of the human linguistic system (Hockett 1960). In this paper, we propose an alternative framework in which, without denying a role for arbitrariness, we propose that iconicity is a powerful vehicle for bridging between language and our experience with objects and events in the world, and as such, we argue that iconicity provides a key to understanding language evolution, learning, and processing (Perniss et al. 2010).

In particular, we argue that iconicity provides scaffolding for the cognitive system to connect communicative form with sensori-motor experience of the world, and that this mechanism is crucial for language development, both in phylogenesis and in ontogenesis, and language processing. In language evolution, iconicity would have been critical in achieving conceptual reference, i.e. reference based on mental representations rather than on purely functional or symptomatic response, and displacement, the core ability of language to refer to things not present in the immediate environment (cf. Bickerton 2009; Kendon 1991). In language learning and vocabulary development, iconicity would play a critical role in supporting the referential mapping process, i.e. learning to map linguistic labels to objects and events in the world. Here, iconic mappings in the language input may provide an additional mechanism for reducing referential ambiguity, helping children to make correct form-meaning associations based on similarities between properties of the communicative form and properties of the referent. Finally, in language processing, iconicity could provide a mechanism to account for the embodiment of language, i.e. the grounding of language in our sensori-motor systems. Under an embodied view, linguistic forms have meaning by virtue of their link with sensori-motor experience of real-world referents. By their very nature of depicting (sensori-motor) properties of referents, iconic mappings imply the engagement of sensori-motor systems in processing the meaning of a linguistic signal.

Iconicity as structure-mapping

Karen Emmorey

In this paper, linguistic and psycholinguistic evidence is presented to support the use of Structure-Mapping Theory (Gentner, 1983; Gentner & Markman, 1997) as a framework for understanding effects of iconicity on sign language grammar and processing.The existence of structured mappings between phonological form and semantic mental representations in sign languages (Taub, 2001) has been shown to explain the nature of metaphor (Meir, 2010) and pronominal anaphora (Schlenker, 2011; Schlenker, Lamberton, Santoro, 2013).Meir's work has shown that the structure of iconic mappings must be taken into account in order to explain possible and impossible metaphorical extensions, which has further ramifications for the possible grammatical devices that can emerge in a language.Schlenker's formal semantic analysis of anaphora reveals explicit effects of iconicity on pronominal interpretation (not available for spoken languages). However, this linguistic work is not framed within a cognitive model, and no claims are made about the role of iconicity for language processing or acquisition.Structure-Mapping Theory proposes a number of psychological principles for relating two representations (e.g., structural alignment, one-to-one mapping, parallel connectivity, non-alignable features), and we suggest this theory can provide a cognitive framework for explaining effects of iconicity on linguistic structure, processing, and acquisition.With respect to processing, we argue that psycholinguistic effects of iconicity may only be observed when the experimental task specifically taps into structured mappings between a phonological representation and a conceptual representation.Further, the nature of the structural alignment between these representations will influence the nature of the processing effects (e.g., whether facilitation, interference, or no effects are observed for iconic forms).With respect to language acquisition, Structure-Mapping Theory predicts that

effects of iconicity will only be observed when the relevant cognitive abilities are in place (e.g.the ability to make comparisons) and when the relevant conceptual knowledge has been acquired (i.e., information key to the iconic mapping).Finally, we suggest that iconicity is better understood as a structured mapping between two mental representations than as a link between linguistic form and human experience.

When the manual modality functions as gesture and/or sign

Susan Goldin-Meadow

The goal of this paper is to widen the lens on language to include the manual modality. We look first at hearing children who are acquiring language from a spoken language model, and find that even before they use speech to communicate, they use gesture (Bates et al. 1979). Moreover, those gestures precede, and predict, the acquisition of structures in speech (Iverson & Goldin-Meadow 2005; Ozcalıskan & Goldin-Meadow 2005, 2009; Rowe et al. 2008). We look next at deaf children whose hearing losses prevent them from using the oral modality, and whose hearing parents have not presented them with a language model in the manual modality. These children fall back on the manual modality to communicate and use gestures, which take on many of the forms and functions of natural language (Goldin-Meadow 2003). These homemade gesture systems constitute the first step in the emergence of manual sign systems that are shared within deaf communities and are full-fledged languages. We end by widening the lens on these sign languages to include gesture, and find that signers not only gesture, but they also use gesture in learning contexts just as speakers do. The findings suggest that what is key in gesture's ability to predict learning is its ability to add a second representational format to communication, rather than a second modality.

Two sources of meaning in infant communication: common ground and act-accompanying characteristics

Ulf Liszkowski

Recent research has shown that infants communicate meaningfully before they have acquired a language. How do they do that? The current paper supports the hypothesis that infants possess social-cognitive skills that run much deeper than language alone, enabling them to understand others and make themselves understood (Tomasello 2008). Recent and new experimental results reveal that infants rely on two sources of extralinguistic information to communicate meaningfully, i.e. to react to and express communicative intentions appropriately. In support, a review of relevant experiments demonstrates, first, that infants use information from preceding shared activities (common ground) to tailor their comprehension and production of communication (Clark & Marshall 1981). Second, a series of novel findings from our lab shows that in the absence of distinguishing information from preceding routines or activities, infants use accompanying characteristics (like prosody and posture) that mark communicative intentions to extract and transmit meaning. Findings reveal that before infants begin to speak they communicate in meaningful ways by binding preceding and simultaneous multisensory information to a communicative act. These skills are not just a precursor to language but an outcome of social-cognitive development and social experience in the first year of life.

Interaction 3: Panel

Self-deselection: The non-verbal accomplishment of "non-next speaker" in face-to-face interaction

Elliott Hoey

Sacks et al. characterize their turn-taking model as one in which turns are " valued, sought, or avoided" (1974:701, emphasis mine). However, most conversation analytic research on speaker-selection has focused on (next) speakers' practices for arranging speakership (see Hayashi, 2013). Thus, the possibility that participants may remove themselves from the set of possible next speakers remains unexplored. I suggest here that practices for displaying no intention to enter the turn space (self-deselection) are indeed available, and are largely non-verbal. The data are 32 instances of self-deselection appearing in video-recorded natural interactions in English; these are analyzed using conversation analytic methods.

Findings indicate that participants' multimodal, instrumental, and vocal resources for self-deselection include gaze aversion, mouth occlusion (e.g., lip closure, hand placed over mouth), sighing, and engagement in unaccountable activities (e.g., self-grooming, eating). These practices occurred in various environments, including, notably, conversational lapses, after changes in participation framework, and at possible topic, sequence, and activity completions. Take, for instance, the following interaction involving three friends, in which a lapse emerges after possible topic completion is achieved via shared laughter.

Orienting to the lapse, Lex reappreciates the sequence with Yeah, occasioning another transition-relevance place. We then observe Marie preemptively removing herself from the set of possible next speakers: in overlap with Lex's Yeah, Marie softly issues the recompleter Hm with closed lips (which precludes most types of talk), then demonstrably seeks a next speaker in turning from Lex to Rachel. Marie's incipient turn toward Rachel is interpretable as possibly selecting her as next speaker (Lerner, 2003), which occasions self-deselection from Rachel: she avoids Marie's gaze while lifting her head up, and settling into an erect posture, she produces a hummed sigh (i.e., voiced with closed lips). These behaviors communicate self-deselection to Marie, who would observe Rachel's gaze avoidance and mouth occlu-

sion, and also audibly telegraph her unwillingness to speak to Lex, who's gazing downward but would nevertheless hear the sigh. Lex finally self-selects, restarting talk. This report contributes to research on multimodality, turn-allocation, and (dis)engagement (Goodwin 1981; Lerner, 2003; Mondada, 2007; Rossano, 2012) by taking an inverse view of turn-taking, wherein participants employ gestural and nonlinguistic resources to coordinate the distribution of conversational negative space. (500 words)

Touching talk: Analyzing the language and social interaction among deafblind persons in tactile Austrian Sign Language

Shimako Iwasaki, Louisa Willoughby, Howard Manns, Meredith Bartlett

This paper investigates how talk and social interaction is organized by Deafblind persons who communicate via a tactile form of Australian sign language (Auslan) in Melbourne. This minority community faces a number of interactional challenges, including the loss of visual and prosodic paralinguistic features in interaction, the modification of a visual language to be received solely through touch, and often quite limited circles of potential conversation partners. Considering that to a large extent human interaction is managed by visual displays such as gestures, facial expressions, and gaze directions as well as hearable resources, it is important to understand how the Deafblind persons who have no access to them manage to initiate, sustain, and coordinate talk in interaction. Thus, the paper analyzes social interaction among Deafblind persons who are fluent tactile Auslan speakers, using the empirical methods of Conversation Analysis and micro-analytic and ethnographic perspectives on multimodality, embodiment, and participation.

This paper explores the interactional practices of Deafblind tactile Auslan speakers, with special attention to the turn-taking organization, overlapping talk, and the use of response tokens produced concurrently during speaker's talk. It focuses on their conduct to see how participants orient to the turn-taking mechanism proposed by Sacks, Schegloff, & Jefferson (1974) in spoken and visual conversations. How turn-taking is organized in signed conversations has intrigued a number of researchers (e.g. Baker, 1977; Baker & Padden, 1978; Coates & Sutton-Spence, 2001; Herreweghe, 2002; Lucas, Bayley, Valli, Rose, & Wulf, 2001; Martines, 1995; among others). However, little work has been conducted in tactle signed conversations among Deafblind persons (except Mesch (2000, 2001, 2013) for Swedish and Berge and Raanes (2013) for Norwegian). Despite the fact that Deafblind conversationalists cannot rely on resoruces such as gaze and non-manual features that are commonly used in sign languages for making meanings and speaker changes, we can observe that turn-taking is managed smoothly in our collection of approximately five hours of data. Rather than adop a discourse analytic appraoch, this study will draw upon the studies of the multimodal and embodied nature of the organization of human action in interaction (e.g. Goodwin, 2003; Goodwin & Goodwin, 2000, 2004; Streeck, Goodwin & LeBaron, 2011) in order to understand the intersubjective worlds of the Deafblind.

Builbing on what researchers of spoken languages have described as the organization of practices for talk-in-interaction, the paper reveals that Deafblind conversationalists reflect conversational strategies managing turn-taking, even if confined to a limited range of modalities. In addition to tactile resources, continuous, but evolving and sequentially- organized information makes it possible for Deafblind participants to sustain coordination. Furthermore, the paper illuminates that they display detailed understanding of talk in progress, which cannot be heard and seen, by performing relevant participation displays at appropriate places. The findings suggest that despite the different range of modalities available, conversation analytic frameworks provide powerful tools for the analysis of talk among Deafblind persons, and we could benefit from the micro-analytic appraoch to have a better appreciation of an intricately coordinated interactional choreography achieved by multiple participants through touching.

Visualising common ground: for communication or cognition?

Mark Tutton, Judith Holler

Common ground (CG), i.e., the knowledge, beliefs and assumptions interlocutors mutually share in interaction, is fundamental to successful communication (Clark, 1996). Next to studies finding gestural ellipsis in the context of CG, an increasing number of studies has shown that speakers use co-speech gestures at the same rate (or even a higher one) when they do compared to when they do not share CG with their interlocutor (e.g. Campisi & Ozyurek, 2013; Holler & Wilkin, 2009; Holler, Tutton & Wilkin, 2011). Common ground (CG), i.e. the knowledge, beliefs and assumptions that interlocutors mutually share in interaction, is fundamental to successful communication (Clark, 1996). In contrast to studies that have found gestural ellipsis when a speaker shares CG with an interlocutor, an increasing number of studies have shown that speakers use co-speech gestures at the same rate (or even higher) when they share CG as opposed to when they do not (e.g. Campisi & Ozyurek, 2013; Holler & Wilkin, 2009; Holler, Tutton & Wilkin, 2011).

There are two alternative explanations for this finding. On the one hand, it has been argued that mentally representing our addressee's knowledge can require considerable cognitive effort (Pickering & Garrod, 2004). In combination with evidence that gesturing helps to reduce cognitive load in cognitively effortful tasks (e.g., Goldin-Meadow, 1999), one hypothesis is that gesture rate is high in CG contexts because cognitive effort involved in mentally representing CG is high. This contrasts markedly with the hypothesis that gesture rate remains high when CG exists because the gestures play an

important communicative role even when they are conveying information that is mutually shared (Holler & Wilkin, 2009). There are two alternative explanations for this finding. On the one hand, it has been argued that mentally representing our addressee's knowledge can require considerable cognitive effort (Pickering & Garrod, 2004). In combination with evidence that gesturing helps to reduce cognitive load in cognitively effortful tasks (e.g. Goldin-Meadow, 1999), one hypothesis is that gesture rate is high in CG contexts because the cognitive effort involved in mentally representing CG is high. This contrasts markedly with the hypothesis that gesture rate remains high when CG exists because gestures play an important communicative role, even when conveying information that is already mutually shared (Holler & Wilkin, 2009).

The present study tested these two hypotheses by combining the manipulation of CG with a manipulation of communicative context. We used a 2(CG) x 3(communication context) between-participants design (18 participants per condition, N=108). All participants watched a short film and narrated it to their addressee. Addressees had either seen parts of the film together with the speaker (CG) or not (no-CG). In addition, we manipulated communication context by asking speakers to narrate their story either face-to-face, via an occluding screen, or into a tape-recorder, a manipulation that has been shown to affect gesture rate in no-CG contexts (Bavelas et al., 2008). If gestures produced in CG contexts are triggered by the cognitive effort of having to mentally represent CG, then social manipulations of this kind should not influence gesture rate in. If gestures conveying information already in CG are communicatively intended, however, then we would expect gesture rate to be different in the three conditions. Our results revealed a significant main effect of social context, with gesture rate being highest in the face-to-face condition, followed by the screen condition, and lowest in the tape-recorder condition. Importantly, we did not find a main effect of common ground on gesture rate, and no interaction between our two factors.The present study tested these two hypotheses by combining the manipulation of CG with a manipulation of communicative context. We used a 2(CG) x 3 (communication context) between-participants design (18 participants per condition, N=108). All participants watched a short film and narrated it to their addressee. Addressees had either seen parts of the film together with the speaker (CG) or not (no-CG). In addition, we manipulated communicative context by asking speakers to narrate their story either face-to-face, via an occluding screen, or into a tape-recorder, a manipulation that has been shown to affect gesture rate in no-CG contexts (Bavelas et al., 2008). If gestures produced in CG contexts are triggered by the cognitive effort of having to mentally represent CG, then social manipulations of this kind should not influence gesture rate. However, if gestures conveying information already in CG are communicatively intended, then we would expect gesture rate to be different in the three conditions.

Our results provide several insights. Firstly, they add to the growing body of evidence for maintained/high gesture rate in some common ground contexts. Secondly, they replicate effects of visual access and dialogue on gesture rate found in earlier studies manipulating social interaction. Thirdly, and most importantly, this social interaction effect affected gesture rates in both the common ground and no-common ground conditions equally. This finding is compatible with the account that gestures representing CG information are communicatively intended but not with a cognitive effort-based explanation. Our results revealed a significant main effect of communicative context, with gesture rate being highest in the face-to-face condition, followed by the screen condition, and lowest in the tape-recorder condition. Importantly, we did not find a main effect of common ground on gesture rate, and no interaction between our two factors.

World-at-hand and World-in-sight: Communication practices in contemporary hetero-technic cooperative work

Eryn Whitworth

This study presents a socially situated account of hetero-technic cooperative work (Reynolds, 1993) using video data of interactions between a doctor and a surgical nurse collaborating on a minimally invasive surgery. Each party controls a distinct technology: the nurse controls a manual instrument used in the production of inter-operative images and the doctor controls a computer that facilitates the annotation and display of inter-operative images. The doctor and nurse's work is reciprocally interdependent (Thompson, 2009) and require each party make step-wise adjustments to their use of technology in light of their collaborator's activity. The configuration of this work environment and technology, and tasks that the doctor and nurse perform confound current understandings of collocated, cooperative work. An analysis of each individual's sensory access to work relevant phenomena, as it is afforded or constrained by a technology, draws investigative attention to a gap in our understanding of embodied communication practices of collocated cooperation. Previous research finds that when individuals work cooperatively in a shared setting they index their immediate environment through co-speech deictic gesture (Streeck, 2009;; Hindmarsh and Heath, 2000) in order to craft a shared understanding of the setting's relevance to their activity (Streeck, 2009). An individual's successful indexical reference of a feature of the shared environment creates a clearing (Dreyfus, 1991) in the shared environment that transforms a previously undifferentiated feature into a mutually recognized focus of interaction. Others have shown that indexical practices are an effective means for collaborators to achieve intersubjectivity in collocated, cooperative work. Indeed it seems that empirical studies of collocated, cooperative work, highlight that environment and context-dependent information made relevant through indexical practices are crucial in a practical ac-

complishment of highly interdependent work. When collaborators survey an environment, without the aid of a technology, or when both collaborators work with equivalent, homogeneous technologies the collaborators have roughly symmetric access to the sensory phenomena that comprise their taskscape (Ingold, 2003) . An individual who references an indexical feature of a shared taskscape leverages her knowledge of her collaborator's symmetrical access to this feature. Her referential practices bring salience to a feature of the environment that is sensorially accessible, but as yet an undifferentiated feature of the surroundings. In doing so the referencing party can add meaning to her speech. In configurations of work where technology induces asymmetries in collaborator's sensory access, as in hetero-technic cooperative work, how do parties achieve intersubjectivity when their worlds-at-hand and worlds-in-sight are disjointed. This study presents a re-examination of two parties indexical reference of a world-at-hand and a world-in-sight in collocated, cooperative work under conditions of technology-induced sensory asymmetries. Examination of hetero- technic work is one step towards embracing a diverse ecology of technologies common to contemporary social interaction, in and outside of professional work.

Cross-Cultural 2: Talks

Typology of gestures and motion metaphors in improvisation of Hindustani vocal music

Tejaswinee Kelkar, Bipin Indurkhya

Singers of Hindustani Classical Music (HCM) are known to teach and sing with elaborate hand gestures displaying melodic contours, rhythmic stresses and other features. These gestures embody a mapping of musical phrases to space, motion and movement. Although the associations between these hand gestures and mental imagery are documented in various ancient and modern treatises, the mapping between the gesture types and their musical content has not been explicitly studied. We explore the chhota-khyal musical form by analysing gestural language in each of its temporal phases. A chhota-khyal is a set of two couplets (sthyayi, antara) involving a fixed set of musical and poetic rules. It can be composed and sung within any raga. We conducted a study of three professional singers of HCM, each performing 5-8 minute khyals in four selected ragas.

The rendition of khyal involves three phases: elaborating the raga by singing on vowels (alaap); tala-bound composition (ciz, bolalap), improvised with theme and variations; and a fast section spanning the whole range of raga (taan, layakari). The transition between different phases of khyal is gradual, and is marked by fixed musical events that a singer must include during the performance. Each musical event in khyal was analysed for all renditions to observe the gestural commonalities pertaining to these events. We analysed the music accompanying gestures in two ways: 1) computer-based motion detection and pitch data extraction, and 2) observational study of palm shapes and their association with motion- metaphors for imaginary sound objects. By analysing pivotal points in khyal across different performers, we propose a typology of gestures used for specific ornamentations in khyal singing. We notated palm shapes and identified gestures in the typology using the schemes described in.

We interviewed the singers regarding their gesturing of musical contours to understand the cognitive construction and awareness-level of these gestures. We observed that semantic content of khyal texts plays a critical role in guiding the mental imagery of the performers. We did a tripartite analysis of these gestures by comparing these results with insights from other controlled experiments on a) prosodic content and gestural stresses in spoken khyal texts; b) abstract presenting gestures in speech; and c) granularity of spatial representations of purely musical melodic contours in listeners.

We compared the spatial coordinates of hand movements with the pitch data for each musical event in khyal. As each phrase is a unique gesture unit, the gesticulation continues throughout the duration of the musical phrase, displaying different components of a gesture phrase. The body only comes to rest once the musical phrase is complete. This neutral position suggests phrase level information in the music, illustrating grouping mechanisms for melodies in HCM to be understood via gestural phrases.

Despite the differences in gestural languages of gharanas, performers and individual rendi- tions, we have illustrated the features of motion metaphors of melody in HCM. The results of this study can be used to generate clearer mappings of intuitive gesture-sound relationships in Hindustani Classical singing, which can be applied to improve pedagogical methods.

Gesture or sign? Hunting languages of the Tsíxa and Ani in Northern Botswana

Susanne Mohr

Recent sign linguistic studies have established a dichotomy of "primary" and "alternate" sign languages (Kendon 2004; Zeshan 2008; Pfau 2012). Primary sign languages are full-fledged systems acquired by deaf people as their L1, while alternate sign languages are "kinesic codes" (Kendon 2004), developed by the hearing members of a speech community for use in special circumstances that preclude vocal communication. Among the latter are for example the Aboriginal sign languages of Australia. Recently, the relevance of hunting languages used among Kalahari Khoe-speaking groups of Southern Africa has been acknowledged in this context (Mohr & Fehn 2012, 2013).

This paper tries to answer the question whether the codes, tshaukakúi among the Tsíxa and lúen among the Ani, belong to the class of alternate sign languages or whether they are

gesture systems that do not display the characteristics of sign languages. The analysis is based on functional as well as structural criteria.

The data consist of video-recordings (approx. 79 minutes) that were collected in spring 2012 in the Kalahari/Okavango region of Botswana. For the elicitation process, a picture/word list of the common Kalahari animals was used.

Functionally, both codes are used on the hunt in order to avoid noises that could scare away prey or attract the attention of predators. The codes are learnt as an L2 by hearing members of the speech communities. In this respect, they are similar to alternate sign languages. A feature that distinguishes them from this sign language type, however, is their relation to the home sign system of deaf Tsíxa in Mababe village. There, four deaf signers use a sign system which is closely related to the Tsíxa hunting code. So far, it could not be determined whether the hunting code serves as a basis for the home signs or vice versa. Nevertheless, the close relation of the two is unique to this speech community.

Structurally, it was established that individual items are phonologically analyzable according to sign linguistic parameters (Mohr & Fehn 2013). However, the handshape inventory of both tshaukakúi and lúen is very small: 16 handshapes in tshaukakúi and 14 in lúen (as compared to 62 in Auslan for example). Moreover, the movement parameter is not applicable to many of the signs. While the symmetry constraint after Battison (1978) can be applied to the codes, the dominance constraint cannot. Further, 52% of the lúen and 97% of tshaukaúi signs are monomorphemic. The few compounds that exist are semantically highly transparent. Morphologically, one piece of evidence points towards a classification as alternate sign languages: the internal ordering of compound elements coincides with the one established by Kendon (1988).

In conclusion, the function of the codes points towards a classification as alternate sign languages. The structural analysis shows clear differences from other (alternate and primary) sign languages, especially for tshaukakúi. In the future, spontaneous data is needed to enable syntactic analyses that provide further insight into the structure of the codes and permit a clearer categorization.

A typology of gestural interaction in Karnatak vocal lessons

Lara Pearson

This paper presents findings from an exploratory study on gestural interaction in South Indian Karnatak vocal lessons. While there has been substantial research on gesture in Indian musical performance (Clayton 2005, 2007; Leante 2009; Rahaim 2012; and Moran 2013), there has been relatively little work on gesture in pedagogic contexts. Although Rahaim (2012) has explored the subject of gesture in music transmission, a systematic analysis of gestural interaction in a complete Indian vocal lesson is yet to be published. The present paper seeks to address this gap in the literature by presenting a case study based on analyses of video recordings showing both teacher and student in one-to-one vocal lessons. The principal aim of the study is to create a typology of gestural interaction in this South Indian pedagogic context.

Video recordings of lessons given by expert vocal teachers based in Karnataka, and Tamil Nadu are analyzed through transcription and coding of the gestures, verbal productions, and musical phrases performed by both teacher and student. The interpretation of this data is based on interviews with participants, and supported by the author's experience as a student of Karnatak music in South India over the past seven years. The findings of the analysis are discussed in relation to typologies of gesture in musical contexts, in particular Simones, Schroeder, and Rodger's (2013) categorization of physical gestures in piano teaching, and Clayton's (2005) typology of gestures in North Indian vocal performance. An assessment is made regarding whether the pre-existing categories presented in these studies are suitable for the analysis of Karnatak vocal lessons.

The wider implications of gestural interactions identified through this analysis are also discussed here. A particular feature of Karnatak music is the extensive use of gestures that cooccur with musical vocalizations. Clayton (2005) classifies similar gestures in North Indian performance as ideographs, and describes them as analogous to melodic flow or motion (p. 376). While in most cases both student and teacher use such gestures, social and ethical pressures control the extent of their production. The effects of these pressures on gestural interaction in this context are considered in the paper. Another feature of Karnatak vocal lessons discussed is the phenomenon of "addressee gesture" (Alibali and Nathan 2012) in which the teacher gestures while the student speaks, or, as in this case, sings. Finally, the advantages of employing sound-accompanying gestures when learning a musical form that relies heavily on improvisation and memorization are considered with reference to studies on gesture and learning, including Goldin-Meadow, Nusbaum, Kelly, and Wagner (2001), and Cook, Duffy, and Fenn (2013).

Grammar and gesture: Panel

Multimodal constructions for dialogue management: on the role of eye gaze and gesture

Geert Brône, Bert Oben, Jelena Vranjes, Kurt Feyaerts

It is a well-known fact that managing the flow of a dialogue in terms of sequentially organized speaker turns is a complex undertaking that requires a high degree of synchronized behavior across interlocutors. Early work in conversation analysis (e.g. Schegloff 1972) already pointed at the intrinsic multimodal nature of dialogue management, but nevertheless, the focus has primarily been on (para)verbal signals guiding the speech exchange system (cf. Schmitt 2005 for

a review). Only recently, the role of other semiotic channels (gesture, posture, gaze) has been studied more systematically. Most notably, there is a growing interest in the emergence of composite signals (Clark 1996, McNeill 2006) involving multiple channels jointly contributing to dialogue management. A key feature is the sequential organization of the different channels, as illustrated e.g. in Mondada (2007) for pointing-gestures as a resource in establishing speakership. The present study pursues this line of research and focuses on the role of eye gaze patterns in relation to other signal systems in dialogue management.

Existing studies on the role of eye gaze in face-to-face interaction, from the early pioneering work by Kendon (1967) to the more recent studies by Vertegaal et al. (2001), Bailly et al. (2010) and Jokinen (2010), have revealed general distributional patterns across speakers and hearers. In this study, we focus more specifically on

(i) microphenomena, or detailed gaze events of about 200-400ms (e.g. brief moment of gaze aversion), associated with specific dialogue acts (e.g. holding the floor);

(ii) the temporal positioning of these gaze events vis- vis other relevant markers in speech and gesture.

An analysis of co-occurrence patterns across different modalities was performed on a multimodal video corpus (InSight Interaction Corpus, authors 2013). This corpus consists of two- and three-party interactions, with head-mounted scene cameras and eye-trackers tracking all participants'visual behavior, providing a unique "speaker-internal" perspective on the conversation. The analysis reveals

(i) specific patterns of gaze distribution related to the temporal organization of dialogue acts. Different dialogue acts typically display specific gaze events at crucial points in time, as e.g. in the case of brief gaze aversion associated with turn-holding, and shared gaze between interlocutors at the critical point of turn-taking;

(ii) a strong correlation and temporal synchronization between eye gaze, speech and gesture in the realization of specific dialogue acts, as shown by means of a series of cross-recurrence analyses for specific turn-holding mechanisms (e.g. verbal hesitation markers co-occurring with gesture and brief gaze aversion) .

The results are indicative of the existence of multimodal discourse constructions, defined along the lines of McNeill (2006)- as recurrent patterns in the physical co-occurrence, functional trade-off and interactive set-up of multiple modes of representation in face-to-face settings.

Existential constructions in gesture: On the grammatical affinity of 'giving' and 'holding'

Irene Mittelberg

Starting from basic object-oriented manual actions of giving and holding, this paper aims to trace crossmodal processes of embodied grammaticalization in English and German. Considering that such routinized physical interactions with the material and social world may be understood as blueprints for prototypical ditransitive and transitive constructions (Goldberg 1995; Bergen & Chang 2005), I propose that they may also be at the root of less transitive existential constructions occurring in both speech and gesture. The entry point to the rationale developed here is the observation that existential constructions in English and German recruit different kinds of verbs. Whereas English there is combines unstressed there (exhibiting presentative, not locative function) with a form of be, German es gibt consists of the non-referential pronoun es ('it') and the lexical Verb geben ('to give') conjugated accordingly. As a well-documented path of grammaticalization evidences, this impersonal usage of geben goes back to the source meaning of manually giving something to someone and related ditransitive constructions involving an agentive subject transferring a physical object to an animate receiver (Lenz 2007; Newman 1998). This could lead us to hypothesize that English existential uses of there is tend to be accompanied by indexical gestures and German es gibt by reduced variants of iconic gestures metonymically alluding to actions of giving (AUTHOR et al., in press). However, the point to be made is that gestures abstracted from physical actions of giving and holding may, even in the absence of the corresponding verbs in the linguistic track of the utterance, fulfill pragmatic functions of physically articulated existentials.

Drawing on English and German multimodal discourse data, it will be demonstrated that although speech- gesture pairings identified in the data confirm the above hypothesis, both existential constructions may be underpinned by variants of the frequently occurring palm-up open hand gesture (M-ller 2004) or by bimanual gestures seemingly handling virtual objects or discourse contents. Instances of such gestural signs with a muted indexical ground were previously found to simply point to the existence of ideas mentioned in speech by providing a tangible surface or a sort of container for the things or ideas alluded to (AUTHOR in press).

These considerations will be backed up by sign language research at the "gesture-language interface" (Wilcox 2004:43) suggesting that when entering the linguistic system, gestures may become lexicalized and grammaticalized in various ways (e.g., Janzen & Shaffer 2002; Kendon 2008). Although gestures tend to be less codified than signed or spoken linguistic signs, gestural existential markers may, as will be shown by way of various examples, exhibit characteristics comparable to those that commonly drive gradual processes of codification and grammaticalization. These may include, among others, reduction in phonetic form, subjectivation (Langacker 2002; Mortelmans 2006); semantic bleaching (Hopper & Traugott 2003); pragmatic/metonymic inferencing (ibid.); ritualization (Haiman 1994); as well as differing degrees of iconicity and schematicity through abstraction (Givón 1985). Overall, the insights presented here shall lay the ground for larger-scale empirical investigations at the intersection of grammaticalization and gesture pragmatics.

English aspect marking constructions in co-speech gesture

Jennifer Hinnell

Discussions of iconicity in co-speech gesture first centered on iconicity in representational gestures (McNeill, 1992), whereas more recent studies have examined iconicity through form-based analyses (Bressem, 2012; Ladewig, 2011; Harrison, 2010). In this multimodal corpus-based study, I examine iconicity in co-speech gesture at a more fine-grained level, using quantitative measures of movement, number of strokes, degree of synchrony of onset, and other parameters. A pilot study (author, 2013) established the consistent marking of aspect in periphrastic auxiliary constructions in co-speech gesture. This study uses a more extensive dataset to establish a "gestural profile" for each auxiliary. It contributes to a growing body of work examining whether, and if so, how, higher order grammatical notions are represented in co-speech gesture.

The dataset consisted of 50 instances each of five aspect-marking periphrastic auxiliary constructions in English continue, keep, start, stop, and quit occurring in their most frequent construction according to COCA. For continue this is the infinitival construction [continue to VVI] (e.g."continue to be"); for the remaining auxiliaries the progressive [aux VVG] construction is most frequent (e.g. "keep going", "start talking"). Data analyses were conducted using the archive and facilities of the Distributed Little Red Hen Lab, an online monitor corpus of 200,000 hours of audiovisual data from public broadcasts, growing at 110 hours/day. Each instance consisted of a video clip and corresponding transcript and was annotated for the presence of gestural co-occurance with intonation unit containing the auxiliary construction. This provided a preliminary measure of gestural. I then annotated each of these positive instances for:

(a) plexity (Talmy, 2000) of the gesture annotated as uniplex/multiplex, and further specified by number of action phases per stroke (Kendon, 2004);

(b) the semantic relation of the gestural phrase to the aspectual auxiliary (measure of iconicity of gesture);

(c) (a)synchronicity of onset between the speech cue and gesture phrase measured in milliseconds.

This study supports the evidence presented in (author 2013) that aspect is reliably and consistently marked in co-speech gesture, and presents a more fine-grained examination of the nature of the iconicity. It is part of a larger study investigating the correlation between linguistic cues, intonation, gesture and body movement, where robust correlations would suggest that the notion of construction (Goldberg 2006; Steen & Turner 2013) should be extended from speech to include embodied components. Further research will include head movement and gaze analysis.

Inflection-specific gestural constructions in English: The case of the catenative auxiliary

Sally Rice, Jennifer Hinnell

For over a decade, an increasing number of linguists have called into question the concepts of the lemma and phrase structure category as the most optimally relevant units for conducting morphosyntactic analysis. Theoretical attention has begun to shift towards constructions based on actual inflected forms of a word. Research in this vein, which is corpus-based and has tended to take a cognitive/functional perspective, includes findings by Bybee & Hopper (2001), Thompson & Hopper (2001), Tao (2003), Stefanowitsch & Gries (2003), Knowles & Zuraidah (2004), Newman & Rice (2004, 2006), and Rice & Newman (2005). Most of these studies are centered around the collocational, constructional, and grammaticalization behavior of inflected forms of individual verbs. Rice and Newman, especially, have posited the idea that inflectional categories and syntactic constructions are "islands" which strand particular lexical items. Under their inflectional island hypothesis, semantic properties tend to inhere in specific morphosyntactic inflections of a lexical item, especially in register-specific contexts. These properties may not spread across all the inflections to characterize the lemma as a whole and so they have proposed the idea of a WIC or word-in-context as the most ecologically veridical starting point for lexico-syntactic research. In this paper, we extend this concept to gestures which accompany speech and suggest that the GIC or gesture-in-context may also be a fruitful focus for corpus-based, construction-centred investigation into linguistic multimodality.

Our starting point is a set of returns from the Little Red Hen media corpus (Steen & Turner 2013), wherein we searched unscripted, interactive video samples for all spoken strings containing the following aspectually charged, auxiliating English "co-verbs": START, STOP, CONTINUE, KEEP, and QUIT, as in I started to leave, you need to stop promising, he keeps doing the same thing, and so on. From all the corpus returns that featured good video of the speaker, we initially coded the frequency for which our target catenative auxiliaries were spoken with an accompanying gesture. Of those hits which inspired a co-speech gesture, we analyzed the gesture for a variety of features, including the asynchrony of the gesture onset with the spoken auxiliary, the complexity of the gesture (number of iterations or separately articulated phases), the relative directionality (towards or away from the speaker), as well as its absolute directionality (along a predominately horizontal or vertical axis). That first-phase analysis yielded statistically significant differences in the gestural profiles accompanying each auxiliating verb. We subsequently have added other tags in the data frames we've constructed for each periphrastic auxiliary verb, including information about the subject, TAM marking, the collocating verb (e.g., LEAVE, PROMISE, and DO, in the above examples),

and other lexico-syntactic properties of the surrounding sentence. Preliminary analysis suggests that certain lexical and grammatical items reliably trigger the use of gesture in conversation, but more importantly, that the incidence and shape of gesture are both affected by inflectional skew, suggesting that co-speech gestures may be semi-constructional in nature and, like speech, highly susceptible to conventionalization and somewhat resistant to characterization at the level of the lemma.

Parallel Session 5 – Wednesday, 9th 4:00 pm – 5:40 pm

Early iconic gesture: Panel

Do comprehension and production of iconic gestures go hand-in-hand?

Leslie Hodges, Şeyda Özçalışkan, Rebecca Williamson

Children frequently gesture at the early ages, and these gestures play a significant role in language acquisition. Importantly, different gesture types appear in children's repertoire at different times, suggesting a unique role for each gesture type in children's spoken language development. The earliest gestures children use, typically beginning around 10 months, are deictics' gestures whose referential meaning is given entirely by the context and not by the form of the gesture (e.g., point at bottle). At this early stage, deictic gestures offer children a tool to refer to objects before they have words for those objects (Iverson & Goldin-Meadow, 2005). In contrast, iconics' gestures that capture the actions or attributes associated with referents in their formXemerge somewhat later, typically several months after children begin to produce their first words (Özcäliškan et al., 2013). Additionally, the majority (74%) of these early iconic gestures convey action information associated with referents, suggesting that children encode action information more readily than attribute information in their early iconic gestures. In this study, we ask whether children's comprehension of iconic gestures mirrors the pattern observed in their production of iconic gestures' with earlier comprehension of iconic gestures conveying action information than those conveying attribute information.

To explore this question, we presented 30 2-, 3- and 4-year-old children's all learning English as their native language's with an iconic gesture comprehension task. Each child was presented with 12 iconic gestures accompanied by neutral speech; half of the gestures conveyed action information (e.g., I have this one + FLAPPING ARMS ON SIDES OF BODY AS IF FLYING ACTION OF A BIRD) and half conveyed attribute information (e.g., "I have this one'+ LINKING HANDS BY THUMBS WITH FINGERS OUTSPREAD AS IF SHAPE OF A BIRD) associated with a referent. After the presentation of each iconic gesture, the experimenter presented the child with pictures of two referents, one matching (correct choice, e.g., bird) and the other not matching (incorrect choice, e.g., basketball) the referent conveyed in gesture, and asked the child to make a choice ("Which one did I have"). Every child saw both the action and the attribute gesture for each referent, and the referents consisted of items familiar to children of this age range. We also controlled for gesture viewpoint by presenting gestures to children with either a character or an observer viewpoint.

Our preliminary results showed that children began to understand iconic gestures by age 3 above chance levels, but only for iconic gestures conveying action information. This initial difference in the comprehension of iconic gestures conveying action vs. attribute information disappeared by age 4, at which point children showed comprehension of both iconic gesture types close to ceiling levels. The analysis indicated no effect of viewpoint on comprehension of either iconic gesture type. Overall, our results suggest that children's comprehension of iconic gestures follow a pattern similar to their production of such gestures, with comprehension emerging earlier for iconic gestures conveying action information.

A cross-cultural examination of frequency and form in two-year-olds' gestures

Paula Marentette, Paola Pettenati, Arianna Bello, Virginia Volterra

Developmental psychologists have explored symbolic play, and in particular children's actions and gestures about objects, as a means of tracking symbolic development. Researchers have focussed on two types of gestures: those that use the hand to represent the object (called body-part-as-object) and those that use the hand as if it were holding the object (called imagined- object). Previous research shows that 3-year-olds produce and comprehend primarily body-part- as-object gestures while 6-8 year-olds prefer to use imagined-object gestures (Boyatzis & Watson, 1993; O'Reilly, 1995; Overton & Jackson, 1973). This study differs from previous studies by examining the gestures spontaneously produced by two-year-olds across two cultures, Italian and Canadian English. We analysed the frequency of gesture production, the range of representational gesture types produced, and cultural differences in type and frequency.

We report on the analysis of the gestures of 21 Italian-speaking children and 42 English- speaking Canadian children produced during a verbal task (Picture Naming Game, PiNG, Bello et al., 2012). Italian children produced twice as many gestures as Canadian children (Italian n=93 gestures, M=4.4 gestures/child; Canadian n=83 gestures, M=2.0 gestures/child) despite being matched by CDI quartile.

Four gesture types were coded, including size-and-shape, body-part-as-object, imagined- object, and own-body. Children did not differ by language in distribution of gesture type. In contrast to the developmental expectation that two-year olds would primarily produce body- part-as-object gestures, children produced similar numbers of imagined-object gestures as body- part-as object gestures. Analysis of the most frequently produced gestures (5 productions, 14 gestures accounting for 75% of data) revealed that 10 of the14 had a preferred gesture type across groups and children (size-and-

shape: small, long; imagined-object: driving, opening, pushing; body-part-as-object: umbrella, comb; own-body: kissing, swimming, washing). Only 4 targets (fork, in, spinning, talking) revealed a mixed pattern, with several gesture types being used for each. Body-part-as-object gestures were equally divided between target images depicting objects and events. Imagined-object gestures were more frequently produced with target images depicting events. These results call into question the assumption that the frequent production of body-part-as-object gestures in young children reveals a deficit in symbolic communication. Instead, the data suggest that the type of gesture produced by young children reflects the nature of the item as well as its presentation and further reveals communicative development.

The analysis of these gesture types has rarely extended to children this young. In this paper we will discuss the implications of the differences in gesture frequency between Italian and Canadian children as well as the similarities in gesture type or form. We will situate this discussion of gesture type in the context of differences in viewpoint as explored in adult gesturers (Brown, 2008; McNeill, 1992; Parrill, 2010) as well as the differences in types of classifiers used in signed languages (Brentari et al., under review; Schembri et al., 2005).

Early Iconic Gesture: Comprehension, Production, and Implications "A cross-cultural examination of frequency and form in two-year-olds' gestures. Preliminary data on Italian deaf signing children (performing the PiNG task in LIS, Italian sign language) shows that these children produce the same four gesture types, and further reveals interesting formational similarities between signs and hearing gestures produced for the same target.

Individual differences in early gesture production predict later cognitive outcomes

Eliza Congdon, Miriam Novack, Dominic Gibson, Susan Goldin-Meadow, Susan Levine

Gesture production is a meaningful predictor of language outcomes in the early years of life For example, children begin to produce deictic gestures, or "points', around 10 months of age, often before they say their first words (Bates, 1976; Masur, 1990) Moreover, individual differences in children's production of deictic gestures are meaningful predictors of various language outcomes (e.g., when a child will say a certain word or produce two-word utterances) (Iverson & Goldin-Meadow, 2005). Although much is known about the rates and consequences of early deictic gestures, less is known about children's early production of iconic or "symbolic" gestures. In one study, children who produced more iconic gestures at 19 months showed larger productive vocabularies at 24 months than their peers (Acredolo & Goodwyn 1988). More recent research has shown that on average, children's iconic gesture production increases from 14-34 months with an iconic gesture "spurt" at 26 months ('oz'oal'o'okan & Goldin-Meadow, 2013). However, no one has yet explored how individual differences in iconic gesture production over this longer time period relate to later language outcomes and other cognitive measures.

We looked at spontaneous gesture production during naturalistic observations of 56 typically developing children interacting with their primary caregiver across a 3 year time period (14-50 months). The children were selected from a larger sample of families who were involved in a longitudinal study of language development. The families were visited in their homes every 4 months and videotaped for 90 minutes. Their speech and gestures were transcribed, and gestures were coded for type (deictic, iconic, conventional). At 54 months, children were tested on a battery of language-related assessments including vocabulary comprehension (PPVT), spatial vocabulary, and a non-verbal spatial relations analogy task.

Individuals varied widely in their total gesture production across the 3 years for all three types of gesture (M_{Iconic}=38.3; $Range_{Iconic}$=3-150, $M_{Deictic}$=752; $Range_{Deictic}$=203-1549, $M_{Conventional}$=752; $Range_{Conventional}$=26-528). Furthermore, amounts of the three gesture types were significantly correlated within children (r=.27, p $<$.05, for iconic and conventional; r=.39, p $<$.01 for iconic and deictic; r=.57, p $<$.01 for deictic and conventional). Despite these strong correlations, the three types of gestures had differing relationships to later language outcomes. While both iconic and deictic gesture production predicted PPVT scores at 54 months (p $<$.05 for both), rates of conventional gesture production were unrelated (p=.46). Furthermore, only iconic gesture was predictive of performance on the spatial analogy task (p $<$.05 for iconic; p=.30 for deictic, p=.986 for conventional). Both deictic and iconic gesture predicted spatial language comprehension (p $<$.05 for both), but when entered into the same regression model, iconic was more predictive than deictic (p=.07 for iconic; p=.23 for deictic). These results provide preliminary evidence that individual differences in early gesture production may be a meaningful way to understand later variation in language skill. Moreover, while rates of iconic and deictic gestures predicted general vocabulary, iconic gesture specifically predicted performance on tasks that required understanding spatial language and relations. Thus, iconic gesture may contribute uniquely to spatial tasks because it is particularly well-suited for capturing and expressing complex spatial information.

Tracing gesture development: Does iconic gesture become more abstract over time?

Erica Cartmill, Miriam Novack, Susan Goldin-Meadow

Children begin to produce iconic gestures later in development than conventional and deictic gesture forms, and their use of iconic gestures undergoes a rapid period of growth around 2.5 years (Özcalıskan & Goldin-Meadow, 2012). Experimental studies have measured comprehension (e.g., Namy et al., 2004) and production (e.g., Overton & Jackson,

1973; Boyatzis & Watson, 1993) of iconic gesture over childhood. Less is known, however, about the development of iconic representation in the gestures children spontaneously produce in natural conversational settings.

Iconic gesture can represent objects, actions, or events in several ways: the hand can represent a hand acting on the world (H=H); the hand can represent an object acting or being acted upon (H=O); or the hand can trace an outline of an object or the path of an action, providing no information about the doer or done-to (H=tracer). H=H gestures (e.g., grasping and pulling the handle of an imaginary drawer) are arguably the most embodied because they closely resemble the movements a person would use to perform an action. No-information gestures (e.g., tracing the path of a flying bird, or tracing the outline of a house) might seem to be the most abstract, because handshape conveys no representational information the iconic elements are carried only in the movement.

If H=tracer gestures are indeed abstract, streamlined types of iconic representation, then we would expect them to emerge late in development, after children have mastered more veridical forms of iconic representation. If, on the other hand, H=tracer gestures are simple forms that are easy to produce because they do not require representation at the level of handshape (i.e., they represent only the path of movement or outline of shape), we would expect them to emerge early, before more complex iconic gestures.

We examined the spontaneous production of these types of representation (H=H, H=O, H=tracer) in children's iconic gestures over time. Using a 700-hour video corpus of 50 English-speaking American children interacting with their parents at home, we coded all iconic gestures produced during 90 minute observation sessions occurring every 4 months from 14 to 54 months of age. We asked whether children's gestures over this period show a developmental trajectory from more veridical (H=H/O), to more abstract (H=tracer), or from simpler (H=tracer) to more complex (H=H/O).

We found that H=tracer gestures emerged late and their use grew quickly. H=H gestures emerged early, grew quickly, and decreased as H=tracer gestures took off. H=O gestures were infrequent but were stable over time. Interestingly, H=tracer gestures conveying action information (i.e., tracing an object's path of motion) emerged at the same time as tracer gestures conveying shape information (i.e., tracing the shape of an object), though the iconic relationship between movement and referent varies between these two gesture types. These results support an interpretation of tracer gestures as abstract rather than simple representations, and suggest a common representational system underlying the use of tracer gestures, regardless of semantic content.

Sign & gesture 2: Talks

The role of gesture in directional verbs in British Sign Language: A corpus-based study

Jordan Fenlon, Adam Schembri, Kearsy Cormier

Directional verbs ("agreement" verbs (Padden, 1983) "indicating" verbs (Liddell, 2000)) in sign languages have been the subject of much debate. These verbs, e.g. GIVE in British Sign Language, can be directed towards locations in space associated with their arguments. Some (e.g., Lillo-Martin & Meier, 2011) have argued this modification is obligatory (at least for object marking), is fundamentally the same as grammatical agreement in spoken languages, and is often accompanied by grammatical non-manual markers such as eye-gaze (Neidle et al., 2000). Others (e.g., Liddell, 2000) propose that this modification is fundamentally different from agreement since it represents a fusion of both linguistic and gestural (specifically pointing) elements. To move the debate forward, more data are needed about how directional verbs are structured and used.

Here we consider linguistic and social factors in the use of directional verbs in the BSL Corpus (Schembri et al., 2013). Directional verbs from the BSL Corpus conversation data (25 participants in 4 cities) were annotated in ELAN for linguistic factors including

(1) whether referents were first established in space before the verb or not,

(2) whether they were spatially modified or not,

(3) participant roles for directional verbs which are modified, e.g. 1st-to-2nd person, 2nd-to-1st, 3rd-to-3rd, etc. and

(4) presence of constructed action (i.e. whether non-manual enactment co-occurs with the manual verb sign or not).

Preliminary results based on 481 tokens reveal that modification of directional verbs occurs for both subject (239/313 = 76%) and object arguments (241/391 = 62%), but is not obligatorily for either. When modification does occur, it is often used without explicit establishment of spatial reference; this is the case for 72% of tokens (175/241) modified for object. Furthermore, 3rd-to-3rd person modification, often assumed to be prototypical, is rare, occurring only 8 times in our data. (Examples of prototypical directional verbs in the literature involve explicit establishment of reference via pronouns/determiners and 3rd-to-3rd person modification (e.g. JOHN POINTa MARY POINTb aASKb "John asked Mary"). Finally, constructed action co-occurred with 70% of tokens (101/144) in which eye-gaze was directed towards the object.

Our results appear to align with the view that directional verbs represent a fusion of linguistic and gestural elements (de Beuzeville et al., 2009; Liddell, 2000). The rate of modification suggests that directionality in BSL is not obligatory which may be attributed to ongoing grammaticalisation processes. Additionally, the frequency with which we observe constructed action with eye gaze suggests a gestural analysis is more appropriate for this phenomenon. We also relate directional modification to social factors including region, age, gender, and language background - something not considered with directional verbs in any sign language before. This is important for understanding variation and language change. These findings highlight the importance of using corpus data

for (sign) linguistics research, to verify or counter previous claims based on little data.

Unlike gestures, production of ASL signs impairs word retrieval for ASL-English bilinguals

Jennie Pyers, Karen Emmorey, Tamar Gollan

Co-speech gesture has been argued to facilitate lexical retrieval, and English speakers who were allowed to gesture had fewer retrieval errors and resolved more tip- of-the-tongue (ToT) incidents than speakers who were restrained from gesturing. Representational gestures seemed to play the strongest role in successful lexical retrieval, indicating that the semantic features represented in the gesture can prime difficult-to- retrieve words. An open question is whether all representational manual movements, including signs from an established sign language, facilitate spoken lexical retrieval. Bimodal bilinguals, fluent in ASL and English, have been found to spontaneously produce ASL signs when struggling to retrieve English words during a picture-naming task designed to elicit ToTs and the production of ASL signs delayed or prevented the resolution of ToTs. Thus ASL signs appeared to block rather than aid lexical retrieval in English. However, such an observation is correlational in nature, and ASL translation equivalents may have been produced as a strategy to aid lexical retrieval rather than reflecting activation of a lexical competitor, resulting in more signs produced for ToTs than for correct word retrievals.

To investigate the role of ASL in resolving English ToTs, we conducted a within- subjects experiment in which 27 hearing native ASL-English bilinguals (M_{age}=32.7, range: 18.03-65.40) performed an English picture naming task, once while their hands were restrained and once while they were allowed to sign or gesture. For the hands-free condition, bilinguals were encouraged (but not required) to produce the ASL translation whenever they had difficulty retrieving the English word. The set of items and the order of conditions were counterbalanced across participants. We observed a marginal facilitative effect of sign/gesture-inhibition on correct retrievals (M_{free} = 40.67, SD=4.02, $M_{inhibit}$=41.59, SD=3.34; t(26)=1.69, p=.10; 95%CI [-2.05; 0.20]), and a significant effect of sign/gesture inhibition on the number of ToTs such that bimodal bilinguals had fewer, as opposed to more, ToT experiences when they were prohibited from signing (M_{free} = 8.93, SD=3.86, $M_{inhibit}$=7.70, SD=3.10; t(26)=2.41, p=.02; 95%CI [0.18; 2.27]). In contrast to gesturing, we found that signing did not facilitate resolution of ToT experiences (MM_{free} = .82, SD=.20, $M_{inhibit}$=.84, SD=.17; t(26)=.59, p=.56; 95%CI [- 0.18;.06]).

Being allowed to sign impaired ASL-English bilinguals' English lexical retrieval, marginally decreasing correct retrievals, and significantly increasing TOTs. We suggest that the negative effect of signs on spoken lexical retrieval differs from the facilitative effect of gestures because signs are lexical items in a bilingual's second language. Both bimodal bilinguals and spoken language (unimodal) bilinguals exhibit more ToTs than monolinguals. For bilinguals, a lexical alternative in one language can compete during lexical selection in the other language, blocking retrieval of the target word. In contrast, gestures are not lexical representations and do not represent a lexical alternative to the target word. Rather, representational gestures can depict an array of features related to the target word, which can serve to prime the semantic features of the target word and thus facilitate retrieval. We are currently conducting an analysis to determine whether the iconicity of the ASL signs impacts ToT rates or resolutions.

Sign language experience changes how learners imitate gestures

Aaron Shield, Richard P. Meier

Physical properties of the visual-gestural modality have important consequences for the articulation of gestures/signs versus vocalizations/words, and for the conditions under which gestures/signs are learned. Because the hands move within a transparent space, children acquiring signs observe them from ever- shifting perspectives. To learn a sign presented in a face-to-face conversation, the child must take the signer's perspective. Failure could lead to production and comprehension errors (Emmorey, 2002; Petitto, 1987).

The acquisition of a signed language offers frequent practice with perspective-taking. Here we explore whether exposure to American Sign Language (ASL) affects how neurotypical adults imitate meaningless gestures. We consider three strategies: (1) mirroring, in which gestures are imitated as a mirror image, (2) perspective-taking, in which gestures are imitated as produced by the model, and (3) learner-centered, in which gestures are imitated as they appear from the learner's perspective. Adults imitated 48 right- handed gestures exhibiting three movement types found in sign languages: vertical (upward, downward), horizontal (inward to or outward from the signer's body), and lateral (left-to-right, right-to-left). We expected subjects to make more errors on lateral gestures, which require perspective-taking; mirroring would lead to errors, such that left-to-right movement would be imitated right-to-left. We did not expect errors on vertical or horizontal items, since these can be correctly imitated by mirroring. Secondly, we predicted that ASL students and expert signers, who have experience with perspective-taking through exposure to ASL, would make fewer lateral errors than non-signers. We also compared subjects' imitation of up/down and inward/outward palm orientations. Mirroring will not produce palm errors, but imitation of inward/outward palm orientations from a learner-centered perspective will yield reversals.

Thirty-four right-handed non-signing undergraduates, 25 fourth- and fifth-semester ASL students, and 18 expert signers (deaf adults or hearing interpreters) participated. We fit a generalized linear mixed-effects model; independent variables were ASL experience and gesture type. Both hypothe-

ses were supported. Subjects were more likely to err on lateral gestures than on vertical or horizontal gestures, $p<.001$. Non- signers had a higher error rate (24%) on lateral gestures than ASL students (14%) or experts (6%), $p<.05$. Although subjects produced palm orientation errors at a low rate, inward/outward palm orientations were no more difficult than up/down palm orientations, even for non-signers.

We next tested another group of 34 right-handed non-signers on these same gestures, but flipped so that the model appeared to gesture with her left hand. We predicted that subjects would no longer make lateral movement errors, since they should be able to mirror movement direction. This hypothesis was supported; subjects made fewer errors on lateral movements in this condition (1%) than in the non-flipped condition (24%), $p<.001$.

Sign-naïve adults frequently use a mirroring strategy in gestural imitation. Some deaf children with autism produce errors suggesting a learner-centered perspective (Shield & Meier, 2012); non-signing adults showed no such tendency. Our results indicate that ASL exposure enables adults to use perspective-taking reliably. We speculate that exposure to sign, but not gesture, promotes perspective-taking.

The simultaneous co-production of Indian Sign Language and spoken Hindi: Syntax and semantics

Ulrike Zeshan

This presentation reports on findings from data involving the simultaneous production of Indian Sign Language (ISL) and spoken Hindi, a behaviour here referred to as "sign-speaking". The sign-speakers who participated in the study work in the bilingual ISL-Hindi environment of a large residential deaf school without formal sign language interpreting provision. When sign-speakers communicate with hearing non-signers and deaf people at the same time, they aim to convey the same information to both hearing and deaf participants despite the syntactic and semantic differences between ISL and Hindi.

Importantly, this phenomenon is qualitatively different from Simultaneous Communication (SimCom) as known from American Sign Language and English co-use. SimCom involves "speaking and simultaneously producing signs' (Liddell 2003:2; see also Tevenal & Villanueva 2009), and where the signing is modified to parallel the structures of English. What is called sign-speaking here, on the other hands, involves "the simultaneous production of two independent and autonomous monolingual utterances' (Donati & Branchini, n.d.:4), a scarcely documented phenomenon that is called "Code-Blending Type 2" in Branchini & Donati (2013). This involves various semantic and syntactic mismatches between the two co-produced languages.

The simultaneous co-use of two languages with different structures raises interesting issues about the nature of sign-and-speech combinations in utterances. One possible outcome is "congruent lexicalisation" (cf. Quinto-Pozos & Adam 2013), where the structures of sign and speech happen to be parallel, so that no conflict exists between structures. However, the ISL-Hindi sign-speaking data also show a range of differences between the simultaneously produced autonomous utterances of ISL and Hindi. These differences can either be syntactic (example 1), semantic (example 2), or both semantic and syntactic (example 3).

(1) SIGN INDEX:1 UNDERSTAND NOT Translation "I didn't understand." Speech mein nahin samjhi 1Sg not understand:PAST.f Translation "I didn't understand."

(2) SIGN INDEX:3 HERE INTERPRETER NOT Translation "For him, there is no interpreter here." Speech unke liye yahan koi nahin hai 3Sg for here any NEG COP:PRS.3Sg Translation "For him, there is nobody here."

(3) SIGN START SELF Translation "(He can) start himself." Speech voh bhi kar sakta hai that also do can:PRS.3mSg COP:PRS.3Sg Translation "He can also do that."

1Sg not The presentation explores the characteristics and potential limits of syntactic mismatches, as well as the distribution of information in cases where the semantic contents of the signed and spoken utterances are different from each other.

Space and viewpoint 1: Talks

Alternating viewpoints: The effects of visual perspective on gestural viewpoint

Fey Parrill, Kashimiri Stec

Previous research suggests that events with a motor action component (e.g., handling an object) tend to evoke gestures from the point of view of a character (CVPT) while events with a path component (moving through space) tend to evoke gestures from the point of view of an observer (OVPT, Parrill, 2010). Events that combine both components (e.g., rowing a boat across a lake) seem to evoke both types of gesture (Parrill, 2010), but it's unclear why a narrator would choose to use one or the other. We ask whether manipulating the visual perspective of the images participants see influences viewpoint in gesture.

35 English-speaking participants read five stories in one of two conditions, Actor or Observer (18 Actor, 17 Observer). Each story consisted of events described by sentences. Each story contained three critical events: For these events, the description also included an image. For participants in the Actor condition, images showed the action from the point of view of an actor (e.g. hands on oars in the prow of the boat; limited view of the lake). For participants in the Observer condition, the same sentence was paired with an image showing the event from the point of view of an observer (e.g. a person rowing a boat in a lake, with each entity in full view). Participants then described these stories.

Gestures co-occurring with descriptions of critical scenes were coded as CVPT or OVPT (or other, but all gestures were either C or O). To control for differences in number of gestures produced by different participants, we compared the mean proportion of CVPT versus OVPT gestures across the two conditions. Participants in the Actor condition were significantly more likely to produce CVPT gestures (mean Actor = .59, SD = .24, mean Observer = .34, SD = .20, t(33) =3.4, p = 0.0018, d = 1.16). (This also means that participants in the Observer condition were significantly more likely to produce OVPT gestures).

Manipulating the visual perspective of the images participants see may influence viewpoint by making the motor action more prominent in the Actor condition and the trajectory more prominent in the Observer condition, which would be consistent with frameworks in which motor action is linked to CVPT gestures (e.g., Hostetter & Alibali, 2008).

The relationship between cognitive skills and co-thought and co-speech gesture frequencies

Lok Chee, Priscilla Shum, Kit Yee Ip, Wing Chee So

Previous research has identified variables that may account for individual differences in gesture production, including spatial and verbal skills (Hostetter & Alibali, 2007), and fluid intelligence (Wartenburger, Kuhn, Sasenberg, Foth, Franz & van der Meer, 2010). A recent study found that representational gesture frequency was negatively correlated with visual and spatial memory, spatial transformation, and conceptualization abilities (Chu, Meyer, Foulkes & Kita, 2013). However, most studies investigated gesture frequency in co-speech conditions only, with few devoted to studying co-thought gestures. The present study examined the factors that account for the individual variations in the gesture frequency in both co-speech and co-thought conditions. Previous research has identified variables that may account for individual differences in gesture production, including spatial and verbal skills (Hostetter & Alibali, 2007), and fluid intelligence (Wartenburger, Kuhn, Sasenberg, Foth, Franz & van der Meer, 2010). A recent study found that representational gesture frequency was negatively correlated with visual and spatial memory, spatial transformation, and spatial conceptualization abilities (Chu, Meyer, Foulkes & Kita, 2013). However, most studies investigated gesture frequency in co-speech conditions only, with few devoted to studying co-thought gestures. The present study examined the factors that account for the individual variations in the gesture frequency in both co-speech and co-thought conditions.

Eighty-seven (39 males) adult participants performed the spatial sequence learning task in which they studied twelve routes, one at a time. There were four routes with ten steps, four with thirteen steps, and four with sixteen steps. After studying a route, approximately half of the participants were assigned to the co-speech condition, and were encouraged to move their hands to rehearse the route while saying aloud the directions from the starting point to the destination once. The remaining participants were placed in the co-thought condition, and were only allowed to gesture without speaking aloud. Before they studied the second route, participants were required to work on a set of mathematics problems for one minute in order to prevent proactive interference. Participants also performed tasks measuring spatial skills (paper-folding task, backward Corsi-Block test), visuospatial memory (Rey-Complex Figure Test- Immediate recall), verbal memory (Hong Kong List Learning TestV First three trials), and fluid intelligence (Ravens). Eighty-seven (39 males) adult participants performed the spatial sequence learning task in which they studied twelve routes, one at a time. There were four routes with ten steps, four with thirteen steps, and four with sixteen steps. After studying a route, approximately half of the participants were assigned to the co-speech condition, and were encouraged to move their hands to rehearse the route while saying aloud the directions from the starting point to the destination once. The remaining participants were placed in the co-thought condition, and were only allowed to gesture without speaking aloud. Before they studied the second route, participants were required to work on a set of mathematics problems for one minute in order to prevent proactive interference. Participants also performed tasks measuring spatial skills (paper-folding task, backward Corsi-Block test), visuospatial memory (Rey-Complex Figure Test-Immediate recall), verbal memory (Hong Kong List Learning Test- First three trials), and fluid intelligence (Ravens).

We counted the number of gestures (e.g., index finger moves to right) and spatial descriptions (e.g., moving to right) produced during rehearsal. Rehearsal duration and recall duration was also recorded. There were no significant differences between the two groups in any of the task scores, nor gesture frequency. Participants in the co-speech condition spent significantly more time rehearsing than those in the co-thought condition, U=461.50, p $<$0.001.The incongruency effect replicates and extends previous work by Kelly et al. (2010) by showing not only a bi-directional influence of gesture and speech, but also of action and speech. In addition, the results show that while actions are easier to process than gestures (Exp. 1), gestures may be more tightly tied to the processing of accompanying speech (Exps. 1 & 2). These results suggest that even though gestures are perceptually less informative than actions, they may be treated as communicatively more informative in relation to the accompanying speech. In this way, the two types of visual information might have different status in language comprehension.

On average, participants in the co-speech condition produced 10.36 (SD=2.77) gestures and 10.41 (SD=2.40) spatial descriptions when rehearsing a route. There were individual variations in the number of gestures (ranging from 3.50 to 18.67) and that of spatial descriptions (ranging from 3.00 to 14.00). Participants in the co-thought condition produced 10.10 gestures (SD=2.93). There were individual variations in the number of gestures (ranging from 3.17 to 14.42).What

were the factor(s) accounting for these individual variations in each condition's In the co-speech condition, we found a positive correlation between spatial skills and the number of spatial descriptions produced, r(41)=.33, p<0.03, and a negative correlation between verbal memory and the number of gestures produced, r(41)=-.31, p<0.05, after controlling for Ravens scores. We counted the number of gestures (e.g., index finger moves to right) and spatial descriptions (e.g., moving to right) produced during rehearsal. Rehearsal duration and recall duration was also recorded. There were no significant differences between the two groups in any of the task scores, nor gesture frequency. Participants in the co-speech condition spent significantly more time rehearsing than those in the co-thought condition, U=461.50, p<0.001.

Gesture frequency was not significantly correlated with any variable in the co-thought gesture condition. On average, participants in the co-speech condition produced 10.36 (SD=2.77; range: 3.50-18.67) gestures and 10.41 (SD=2.40; range: 3.00-14.00) spatial descriptions when rehearsing a route. Participants in the co-thought condition produced 10.10 gestures (SD=2.93; range: 3.17-14.42). What were the factor(s) accounting for these individual variations in each condition's Regression analysis was carried out separately for the two conditions, with gesture frequency as the dependent variable, Ravens as a control variable, and Corsi Block, paper-folding, and HKLLT scores as predictors. In the co-speech condition, we found that verbal learning scores was a significant predictor for number of gestures produced,]=-.34, t(39)=-2.27, p=0.029. The number of spatial descriptions participants produced was predicted by their scores in the paper-folding task,]=.326, t(39)=2.08, p=0.045.

Our results suggest that co-thought and co-speech gestures may function differently. Specifically, when speaking, participants with poorer verbal memory tend to produce more gestures to lighten the memory load during spatial information encoding. However, this function of gesture is not found in the co-thought gesture condition. However, gesture frequency was not predicted by any variable in the co-thought gesture condition.

The beneficial roles of gesture vs. spatial language in spatial recall

Wing Chee So, Kit Yee Ip, Lok Chee, Priscilla Shum

When we navigate in an environment, we form a spatial representation regarding the relation between locations, objects, and paths. However, little research has investigated whether gestures can improve spatial relation learning and memory. A recent research suggested that gesturing while thinking silently (co-thought gestures) is an effective embodied technique in enhancing recall accuracy of steps in a navigation task (So, Vong, Ching, 2012). However, it is uncharted whether gestures accompanying spatial language (co-speech gestures) are as effective as co-thought gestures in facilitating spatial recall. It is also unknown whether spatial language is as effective as co-thought and co-speech gestures in improving memory for spatial relation. Finally, it is not clear whether the facilitating effects of both kinds of gestures and spatial language are found in various levels of task complexity. In our experiments, adult participants performed the spatial sequence learning task in which they studied twelve routes, one at a time. There were four routes with ten steps, four with thirteen steps, and four with sixteen steps. Participants were randomly assigned to one of the four conditions. In the Gesture+speech (co-speech gestures) condition, after studying a route, participants were encouraged to move their hands whenever they think it might be helpful to rehearse the route while saying aloud the directions from the starting point to the destination once. In the Gesture-only (co-thought gestures) condition, participants were encouraged to move their hands but they were not allowed to use speech to rehearse the route. In the Speech-only condition, participants said aloud the directions while holding a softball with both hands. In the no-rehearsal condition, participants read alphabets aloud for 20 seconds in order to prevent them from mentally rehearsing the route. All participants later reconstructed the route with sticks during recall. Before they studied the second route, participants were required to work on a set of mathematics problems for one minute in order to prevent proactive interference. We coded the recall accuracy of each step. A step was considered to be correctly recalled if the placement of the corresponding stick matched the direction and sequence of step in the diagram. A step was considered to be incorrectly recalled if either the direction or order of the stick was not correctly placed. As expected, participants were more likely to recall steps accurately in the less complex routes (10- and 13-step) than in the complex routes (16-step). Interestingly, the proportion of steps recalled accurately was higher in the Gesture-only condition than in the other conditions. Such differences were found in all three types of routes. Participants in the Gesture+speech condition performed slightly better than those in the Speech-only condition. The proportion of steps accurately recalled in the Speech-only condition was comparable to that in the no-rehearsal condition. Our findings suggest that co-thought gesture produces a stronger facilitating effect on learning spatial relation than co-speech gesture. Both kinds of gestures, however, are more effective than spatial language in encoding spatial relation.

Experimental 2: Talks

Recalling the right words: Handedness effects on recalling action verbs learned with action pictures

Jacqueline de Nooijer, Tamara van Gog, Fred Paas, Rolf Zwaan

In this contribution to the thematic panel on Gestures in Learning with Static and Dynamic Visualizations, we present studies investigating whether observing and/or making ges-

tures when learning words with static and dynamic visualizations, can positively affect learning outcomes.

The grounded cognition framework states that sensorimotor experiences play a functional role in knowledge representations (Barsalou, 2008) This suggests that there could be a cognitive benefit to conducting bodily movements or gestures while learning new information In fact, considerable evidence suggests that producing or observing congruent gestures with speech can improve problem solving and language learning (see Goldin-Meadow & Alibali, 2013; Kelly, Manning, & Rodak, 2008). For example, it has already been demonstrated that observing and imitating gestures can aid second language (L2) word learning (see Macedonia, M'oller, & Friederici, 2011; Schmidt-Kassow, Kulka, Gunter, Rothermich, & Kotz, 2010). Our research extends these findings by exploring and comparing the use of gesture, movement and the observation of movement while children learn new words. Furthermore, while many of the related studies on word learning focus on L2, the current study looks at word learning in L1

Across two experiments we compare word learning in 7-8 year old children First we look at word learning when the children simply observe a dynamic animation containing either meaning congruent or incongruent movements In this case, the children did not make any movements themselves In a second experiment, we look at word learning when the words are paired either with: 1) Observing a congruent dynamic animation, but making no movement themselves; 2) Observing a congruent dynamic animation and performing a gesture related to the dynamic animation or 3) observing a congruent dynamic animation and performing an unrelated, but goal-directed movement

Overall our results suggest that simply observing movement as well as observing movement and performing congruent gesture aids word learning Combining the previous results on congruent gesture or physical engagement on cognitive tasks and our current results, we propose that the improvement in learning comes from the opportunity for a person to make a congruent link from the to-be-learned information to their own experiences via the sensorimotor system Gesture is one optimal manner to make this link, but the observation of dynamic movement can also make this link.

Gesture, movement and observation of movement: Evidence from two word learning studies

Lea Hald, Marianne van den Hurk, Harold Bekkering

This proposal is part of a thematic panel on Gestures in Learning with Static and Dynamic Visualizations. Pictures are often used to aid word learning. When learning action words, the Body Specificity Hypothesis states that reading an action word leads to creating body-specific mental simulations of the denoted action (Casasanto, 2009). Left and right-handers should, therefore, make different mental simulations. A recent study investigated effects of seeing a picture with a left-handed, right-handed or bimanual perspective, which either matches or mismatches the participants' mental simulation, on learning object-manipulation words in an artificial language. Results showed that right-handers recalled fewer word definitions, when the pictures seen during learning mismatched the right-handed mental simulation evoked by the verbal definition (De Nooijer, Van Gog, Paas, & Zwaan, in press).

The current study follows up on this finding, by investigating whether a mismatch in mental simulations evoked by a verb and a picture, is only detrimental in a learning task, or whether right-handers would show the same effect in a recall task, which does not involve coupling new words to an existing concept. We presented participants with object-manipulation verbs (e.g., to cut) that were coupled to a left-handed or right-handed perspective picture that reflected the action. The picture-word pairs either followed each other in rapid succession (No break), were presented with a two second break in between (Break), or were presented after the participant made a non-meaningful right arm movement (Movement). After 20 items, a free recall task was presented.

We hypothesized that words coupled to a right-handed perspective picture are better recalled when a right-handed arm movement is made, given the close link between action and language that is proposed by theories of embodied cognition (e.g., Barsalou, 1999) and the finding that arm activation results in faster recognition of action words involving arm movements (Pulverm'oller, Hauk, Nikulin, & Illmoniemi, 2005). This might also facilitate recall. Furthermore, when there is a two second break between items, in which the picture-word pair is not present, the participants might try to memorize the words via the more conscious process of mental imagery, which might lead to a disadvantage for the words coupled to the left-handed perspective pictures, given that the mental representations are incongruent with these pictures for right-handers.

Results showed a significant interaction between condition and hand perspective of the picture. There were no significant differences in the No break and Movement condition. However, in the Break condition right-handers recalled significantly more words when presented with a right- versus a left-handed perspective picture. It seems, therefore, that in the Movement condition attention moves away from the mental simulation to the arm movement, while in the two seconds in the Break condition the focus on the mental simulation can hinder memory when the mental simulation is incongruent with the picture.

We are currently conducting a similar experiment with use of an eye-tracker to investigate whether the effect we found can be ascribed to different looking patterns to the left and right-handed pictures.

Effects of pointing gestures on source memory in young and older adults

Kim Ouwehand

Research has shown that memory for content-context associations (i.e., source memory), is more error prone than memory for content information in isolation (i.e., item memory). In addition, whereas source memory has been shown to suffer from age-related declines, due to problems during information encoding, item memory remains relatively unaffected by increasing age. Enactment studies show that pantomiming action phrases enhances source memory for e.g. action-object associations more than listening to these phrases in both young and older adults. An interesting question is whether gesturing could also assist source memory performance for location-object associations. We conducted two experiments to investigate whether producing pointing gestures at the locations of pictures during encoding would have a more beneficial effect on source memory for picture-location associations than naming the location of the pictures. The main hypothesis was that pointing would lead to better source memory than naming in both young and older adults, but that the difference would be more pronounced in older adults. In Experiment 1, 24 young adults were presented with a source memory task consisting of an encoding and retrieval phase. In the encoding phase participants saw two types of pictures (artificial objects, e.g. "a trash bin" and natural objects, e.g. "a tree") that were presented sequentially in one of the quadrants of the computer screen. A within-subjects design was used in which participants had to point to one type of pictures (either artificial or natural) and name the location of the other type of pictures (either natural or artificial, respectively). In the retrieval phase, participants were presented with the pictures they had seen in the encoding phase, intermixed with new pictures. All pictures were presented in the center of the screen. Participants had to judge whether or not they recognized the pictures from the encoding phase, and to indicate at which location of the screen they had seen the pictures they recognized. The results partly confirmed the hypothesis, showing that source memory performance in the pointing condition was better than in the naming condition. Experiment 2 used the same design and procedure, but this time young adults (n=39) were compared to older adults (n=40). Similar to the results of the first experiment, pointing led to better source memory than naming. In addition, item memory was better in the pointing condition than in the naming condition in both age groups. Young adults outperformed older adults on the source memory but not the item memory task. In conclusion, the results of both experiments suggest that manually pointing at pictures during the encoding phase might be an effective way to improve source memory which seems especially important for older adults as normal aging is associated with reduced source memory performance. In addition, the findings from Experiment 2 suggest a positive effect of gesturing on item memory for both age groups.

The impact of depictive gestures on listener comprehension is linked to working memory

Ying Choon Wu, Seana Coulson

Existing research has yielded conflicting views on the question of how depictive co-speech gestures affect listener comprehension. On the one hand, some studies suggest that when speech and gestures express conflicting meanings, comprehension is impaired, whereas no benefit obtains in the reciprocal case V when speech and gestures are congruent. On the other hand, however, the opposite pattern of outcomes has also been reported V that is, congruent speech and gestures have been shown to enhance word and message- level comprehension relative to a neutral baseline in which gestures are not visible.

The present study explores the hypothesis that the impact of depictive gestures depends in part on individuals' abilities to maintain gestural representations in immediate memory during the dynamic process of multi-modal integration. 84 healthy adults viewed short video segments extracted from continuous footage of discourse in which objects and events were described both through speech and gestures. Three types of videos were presented. On congruent trials, audio and video portions of each clip were paired in their original form. On incongruent trials, audio and video elements were swapped such that substantially fewer semantic mappings between speech and gestures could be discerned. On neutral trials, video segments were replaced with freeze frames (depicting the speaker) extracted from portions of the discourse stream when the speaker was not gesturing. Neutral stills were combined with the same audio segments used in the experimental conditions. Each video prime was followed by either a semantically related or unrelated picture probe. Reaction times and accuracy were collected as participants made relatedness judgments to probes. Additionally, each participant's working memory (WM) abilities were assessed through a novel psychometric instrument dubbed the movement span task (Wu & Coulson, accepted), which measures short-term storage capacity of kinesthetic phenomena, such as body configurations and patterns of movement.

On average, participants classified related pictures preceded by congruent speech and gestures more quickly and accurately than the same items preceded by speech without gestures, suggesting that iconic co- speech gestures can benefit comprehension. Notably, relative to neutral baseline response times, probes were classified more slowly on incongruent trials, suggesting that gestures can also negatively impact comprehension when semantic mappings to the ongoing speech cannot be readily established. The degree to which congruent gestures facilitated picture probe classification speed increased linearly with movement span scores. Moreover, when sorted by movement span scores, it was found that participants with superior WM abilities exhibited primarily facilitation by congruent gestures, whereas those with smaller spans

primarily experienced interference when speech and gestures did not cohere. In keeping with an embodied view of higher order cognition, these findings reveal that cognitive systems important for remembering and reproducing body movement are also related to the process whereby conceptual representations are constructed from the observation of movement produced in the context of talking. They also indicate that listener comprehension is impacted differently by iconic gestures depending on WM abilities.

Parallel Session 6 – Thursday, 10th 10:20 am - 12:00 pm

Action for learning: Panel

Action, gesture and abstraction in mathematical learning

Andrea Donovan, Rebecca Boncoddo, Caroline Williams, Candace Walkington, Elizabeth Pier, Jessica Waala

The embodied cognition framework (e.g., Glenberg, 2010) holds that cognition is grounded in action. This perspective implies that the actions people produce during learning may influence what they learn. Those actions may also influence the gestures they make when later speaking about their knowledge. According to the Gesture as Simulated Action hypothesis (Hostetter & Alibali, 2008), gestures derive from action simulations that underlie thinking and speaking. Thus, learners' actions may leave a legacy in their gestures. Moreover, gestures are themselves a form of action, and as such, gestures may directly influence learning.

In this research, we address three questions about relations between action, gesture and learning in mathematics: (1) Do directed actions affect learning (2) Do the gestures learners produce when speaking about their knowledge reflect the actions they produced when learning. (3) Are variations in learners' gestures associated with variations in learning. Study 1 addresses these issues among children learning about mathematical equivalence using manipulatives that afford different actions. Study 2 addresses these issues among undergraduates directed to perform actions related to and then prove two mathematical conjectures.

In Study 1, children (N = 80) were assigned to one of four conditions to learn about mathematical equivalence. In three conditions, children were guided to model equations with manipulatives (a pan balance, towers of blocks, or buckets with beanbags). In the control condition, no manipulatives were used. At posttest, children in the control condition outperformed those in the manipulatives conditions. However, children in the manipulatives conditions showed greater relational understanding of the equal sign. Further, patterns of gesture production differed across conditions. Children in the buckets condition were more likely to produce two-handed gestures; thus, their actions during learning left a legacy in their gestures.

In Study 2, undergraduates (N = 120) were asked to perform actions either relevant or irrelevant to solving two conjectures that they then attempted to prove. One was a geometry conjecture (the triangle inequality theorem) and one was a mechanical conjecture (about gear systems). Participants provided think-aloud protocols, and their gestures were coded as dynamic or static. Static gestures represent a stationary object, whereas dynamic gestures depict an object that is moving or changing (e.g., a triangle being "stretched" or a gear turning). Participants who were directed to produce relevant actions were more likely than those that produced irrelevant action to generate key insights needed for sophisticated proofs. The likelihood of producing dynamic gestures did not differ for participants producing relevant or irrelevant actions. However, use of dynamic gestures predicted participants' success at generating sophisticated proofs, beyond predictions based on speech alone.

Taken together, these findings suggest that directed actions influence abstraction and learning, sometimes in unexpected ways. Further, learners' actions may leave a legacy in their gestures, by influencing the nature of their action simulations. Finally, variations in gesture are associated with differences in learning. These findings raise questions about the causal pathways that link action, gesture and learning, and suggest that action affects learning via its effect on gestures.

Action for learning, but gesture for generalization: How abstract and concrete movements help children

Miriam Novack, Eliza Congdon, Neon Brooks, Susan Levine, Susan Goldin-Meadow

Previous research suggests that that performing actions affects our internal representations (e.g., James, 2010; Kontra, Beilock & Goldin-Meadow, 2012; Sommerville & Woodward, 2010; Wilson, 2002). Gestures, which are a special kind of action, can also have profound affects on thinking and learning (see Goldin-Meadow, 2003). Here, we bridge these two areas of research through a framework that views action and gesture as points along a continuum of concrete to abstract movement. Actions can be thought of as "concrete" since they interact directly with the physical world, whereas gestures can be thought of as "abstract", since their form and representational content can be similar to actions, yet they do not permanently change the external world. Here, we explore how movements that varied in their "concreteness" may affect learning and generalization across the domains of math and mental rotation in children.

In experiment 1, we taught third-grade children (n=69) a strategy for solving mathematical equivalence problems that was instantiated in one of three ways: (1) a physical action performed on objects, (2) a concrete gesture miming that action, or (3) an abstract gesture that abstractly represented the strategy. We found that all types of training were equally effective for teaching children how to solve the problems on which they were trained. However, only children in the two gesture training conditions succeeded with solving a set of near-transfer problems (p's$<$.05), implying that they had developed a more flexible understanding of the strategy. Finally,

when given far-transfer problems, only children who learned with the abstract gesture were successful ($p<.01$).

The results of Experiment 1 suggest that the more abstract the training strategy became, the more useful it was for generalization. In Experiment 2 we asked whether the benefit for abstract forms of action would hold in a mental rotation task, where the actions being represented were more directly applicable to the problems to be solved. Additionally, we extended our "Action to Abstraction" continuum to include an even more abstract, "imagine action" condition. In training, children were given practice either physically rotating an object, gesturing the rotation of an object, or imaging the rotation of an object. Following training, children were tested on four types of mental rotation tasksV a task that mirrored what children were asked to do during training; near transfer tasks that changed either the type of stimuli or the number of items they were asked to select, or a far-transfer task that changed both the stimuli and task requirements. Preliminary results (n=30) suggest that all three types of training helped children improve on the task they were trained on, but only the most abstract form of movement, imagining rotation, helped children improve on the far-transfer task ($p<.05$). For near-transfer tasks, preliminary trends suggest that both the gesture and imagine conditions, but not action, lead to learning. These preliminary results converge with those of study 1 suggesting that while physical experience acting may help children learn, abstract representations can facilitate children's generalize to new contexts and problem types.

The integration of gestures and actions with speech: Should we welcome the empty-handed to language comprehension?

Spencer Kelly, Meghan Healey, Aslı Özyü rek, Judith Holler

Background: Gesture and speech are theorized to form a single integrated system of meaning during language production (McNeill, 1992), and evidence is mounting that this integration applies to language comprehension as well (Kelly, Ozyurek & Maris, 2010). However, it is unknown whether gesture is uniquely integrated with speech or is processed like any other manual action. To explore this issue, we compared the extent to which speech is integrated with hand gestures versus actual actions on objects during comprehension.

Method: The present study employed a priming paradigm in two experiments. In Experiment 1, subjects watched multimodal videos that presented auditory (words) and visual (gestures and actions on objects) information. Half the subjects related the audio information to a written prime presented before the video, and the other half related the visual information to the written prime. For half of the multimodal video stimuli, the audio and visual information was congruent, and for the other half, incongruent. The task was to press one button if the written prime was the same as the visual (31 subjects) or audio (31 subjects) information in the target video or another button if different. RT and accuracy were recorded.

Results: In Experiment 2, we reversed the priming sequence with a different set of 18 subjects. Now the video became the prime and the written verb followed as the target, but the task was the same with one differenceXto indicate whether the written target was related or unrelated to only the audio information (speech) in preceding video prime. ERPs were recorded to the written targets.

In Experiment 1, subjects in both the audio and visual target tasks were less accurate when processing stimuli in which gestures and actions were incongruent versus congruent with speech, $F(1, 60) = 22.90$, $p<.001$, but this effect was less prominent for speech-action than for speech-gesture stimuli. However, subjects were more accurate when identifying actions versus gestures, $F(1, 60) = 8.03$, $p = .006$.

In Experiment 2, there were two early ERP effects. When primed with gesture, incongruent primes produced a larger P1, $t(17) = 3.75$, $p = 0.002$, and P2, $t(17) = 3.02$, $p = 0.008$, to the target words than the congruent condition in the grand-averaged ERPs (reflecting early perceptual and attentional processes). However, there were no significant differences between congruent and incongruent conditions when primed with action.

Discussion: The incongruency effect replicates and extends previous work by Kelly et al. (2010) by showing not only a bi-directional influence of gesture and speech, but also of action and speech. In addition, the results show that while actions are easier to process than gestures (Exp. 1), gestures may be more tightly tied to the processing of accompanying speech (Exps. 1 & 2). These results suggest that even though gestures are perceptually less informative than actions, they may be treated as communicatively more informative in relation to the accompanying speech. In this way, the two types of visual information might have different status in language comprehension.

Understanding the neural effects of learning with gesture

Elizabeth Wakefield, Eliza Congdon, Miriam Novack, Susan Goldin-Meadow, Karin H. James

Transitive action and co-speech gesture are two forms of movement we use pervasively in daily life. Transitive actions are used to directly manipulate our environment; gestures accompany our speech as we manipulate ideas and communicate with others. We know that both types of movement profoundly affect cognition, and that similar neural networks are recruited when adults perceive actions and gestures. In the current study we investigate how learning through gesture affects subsequent processing of learned information, and consider these findings in relation to the action-learning neuroimaging literature. In other words, we ask whether gesture has its effects on learning because it is grounded in action.

To investigate how gesture use during learning changes how learned information is processed, we conducted a two part behavioral and imaging study. In Part 1, we taught chil-

dren to solve equal-addend mathematical equivalence problems such as 3+4+6=?+6 (Cook et al., 2008; Goldin-Meadow et al., 2009). Half of the participants learned to produce an equalizer speech strategy ("I want to make one side equal to the other side") accompanied by a matching gesture highlighting the two sides of the equation. The other half simply learned to produce the speech strategy alone. After training, children took a written posttest, and those who answered 3 to 6 problems correctly were invited back for a functional Magnetic Resonance Imaging (fMRI) session a week later (Part 2). During Part 2, children were scanned while performing three types of tasks. Children (1) listened passively to the speech strategy while viewing equivalence problems, (2) produced movements while viewing problems (for the Speech-Gesture group, the movement was the learned gesture; for the Speech group it was a novel bilateral tapping movement), and (3) solved additional equivalence problems.

Our Part 1 behavioral results show that children in both conditions successfully learned to solve the equivalence problems. Because we were interested not in whether children could learn, but how they had learned, we limited our fMRI study to children who correctly answered 3-6 questions on the original posttest. We thus asked whether children processed information related to mathematical equivalence (i.e., seeing math problems, hearing the speech strategy, making movement) differently, having learned via different types of training. We asked whether the networks uniquely activated by children who learned through gesture were similar to those previously found to be activated after children learned through transitive action (e.g., James & Swain, 2010). Preliminary results (N = 20) suggest that for children who learned via gesture, solving equivalence problems recruits regions that are activated when children view objects that they learned about via action. This pattern was not seen for children who learned only through speech. We also found that children who learned through gesture recruited a set of overlapping motor regions for processing the speech strategy and producing their gesture, a pattern again not found for children in the speech alone condition. Our results offer promising preliminary evidence for similarities between how learning through gesture or action affect neural processing of learned information.

The intelligibility of gesture as a situated accomplishment: Panel

Gesture in the Seattle DeafBlind Community: a triumphant story about the regeneration of obviousness

Terra Edwards

This paper examines the coordination of gestural and linguistic resources among DeafBlind people in Seattle. Most members of this community were born deaf and due to a genetic condition, slowly lose their vision. They came to Seattle using Visual American Sign Language (VASL) and as vision deteriorated, they learned to receive VASL signs tactually. However many gestural and linguistic cues are produced by the face and other areas of the signer's body. DeafBlind people only had tactile access to the hands of the signer. Therefore, these facial and bodily cues became inaccessible. Over time, this led to a deterioration of sighted intuitions regarding the meanings of gestures. Things like a shrug, raised eyebrows, hunched shoulders, and pocketed hands, became cryptic, surprising, and even disturbing when encountered. Then, in 2007, the "pro-tactile movement" was initiated and collective efforts shifted from developing compensatory strategies for coping with vision loss, to establishing new, tactile modes of interaction and communication. This changed how DeafBlind people were orienting to signs, objects, and other people in the immediate environment, and eventually led to the emergence of new grammatical sub-systems in Tactile American Sign language (TASL). Drawing on 18 months of ethnographic fieldwork, more than 190 hours of videorecordings of interaction and language use, and 50 interviews with members of the community, this paper examines interactional and grammatical processes through which bodily movements accrue "transparent" and "obvious" meanings among DeafBlind people. I argue that the notion of iconicity is often bound up with problematic assumptions regarding transparency, and I propose an alternative conceptual approach for understanding how gestural obviousness is regenerated by DeafBlind people in interaction.

The interactive work of rendering a gesture intelligible

Charles Goodwin

In much gesture research the meaningfulness of gesture is taken for granted, and assumed to be transparent: "Human beings K find such gestures as pointing and pantomiming totally natural and transparent: just look where I am pointing and you will see what I mean" (Tomasello, 2008). Gestures produced by a man with severe aphasia- his vocabulary is restricted to Yes, No, and And- provide an opportunity to strongly question such an assumption, and to systematically investigate the seen but unnoticed interactive practices that render gesture intelligible. In the sequence that will be the primary focus of this presentation Chil, while talking with two of his adult children responds to a spate of their talk by using his arm and hand to enact a multi-part pantomime. His addressees greet this gesture with blank uncomprehending stares. They begin their work to provide for its intelligibility by trying to link it to the talk that occurred just before it. That prior talk depicts a number of different events that might constitute a point of departure for how Chil's gesture is to be understood, and initially an addressee picks the wrong one. Tomasello (Tomasello, 2008) argues that shared common ground provides for the transparent meaningfulness of a

gesture. It would be difficult to find participants who shared more common ground in terms of both shared life history and immediately prior talk than these, but that shared history is insufficient to establish the meaningfulness of Chil's gesture. The participants establish what Chil's gesture was saying by first using talk to place candidate proposals about what the gesture might mean in a public environment, and then progressively modifying that structure by changing parts of it. Chil participates in this process both by accepting or rejecting their proposals, and through the production of new pantomime and pointing gestures. His initial gesture is rendered intelligible through the systematic modification of unfolding action within a co-operative transformation zone in which multiple participants contribute to an emerging structure of intelligibility by progressively performing systematic transformative operations on materials produced by others. Though Chil's inability to explicate his gesture(s) with rich talk might seem to constitute a special, atypical case, it will be argued that gesture more generally is seen as transparently meaningful through the way in which it emerges within specific arrangements of co-occurring resources that are being progressively and systematically changed to provide for the local intelligibility of emerging action, that is to co-operatively establish just how a moving hand is to be understood at this particular moment. The intelligibility of gesture is build upon a dense network of public, interactive practice.

Teachers and students' collaborative work to render pointing gestures intelligible

David DeLiema

Gesture is a central resource for shaping the knowledge of newcomers in complex environments such as a geology field school. How do teachers and students maximize the intelligibility of such gestures. In this presentation, I use video recordings of a geology field school in Yellowstone National Park to examine the resources teachers and students recruit in interaction to help render intelligible deictic gestures. In thinking of deictic gestures as hand movements that draw attention to objects in the environment, I ask: What other actions surrounding the hand help interlocutors to make the "point" able to be understood. Intelligibility, in this way, involves at once crafting resources that increase the clarity of the gesture and ensuring that the gesture has been understood.

I examine a natural experiment that unfolds at the field site: The teacher works to explain the same geology structure, a "rock fold," to five different groups of students over 22 minutes. The first group is a cluster of students scattered far away from the fold and the remaining four groups are positioned close to the fold in the following configurations: 3 students, 5 students, 2 students, and 1 student. Across these five interactions, the students and the teacher work with each other in careful coordination with the environment to increase the intelligibility of their deictic gestures. In all, interactants draw on 13 publicly observable resources.

Vision

1. Visual inspection of referent prior to and during the point (e.g. looking around while saying, "where did I see that")
2. Visual inspection of other people in the interactive field (e.g. glancing at listener)

Body

3. Variation in gesture form (e.g. flat hand, single finger point, two-hand point)
4. Physical contact with referent (e.g. stepping on referent, touching referent)
5. Body movement transporting the hand horizontally (e.g. pointing while swinging the arm; pointing while walking)
6. Body movement transporting the hand vertically (e.g. bending down while pointing)
7. Body movements that clear the pathway between the gesture, the referent, and the speaker/listener (e.g. moving one's torso to make the gesture and the referent visible)
8. Listener-gestures about referent (e.g. student points toward referent while professor talks)

Talk and body

9. Presentation of categorical/conceptual symbol for the referent (e.g. talking about a "fold" or gesturing a symbol for fold)
10. Depiction of visual appearance of referent (e.g. talking about a fold closing and coming around... or gesturing a sharp angle)
11. Description of external objects populating space around referent (e.g. pointing to rock cairn; describing the fold as "where your heels are...")
12. Epistemic markers (e.g. "okay," "I know...," or nodding or raising eyebrows)

Talk

13. Deictic linguistic terms (e.g. here, there, that, this)

Multiple examples of each resource will be shown in the presentation. This list describes the publicly observable resources on which participants draw to enhance the intelligibility of deictic gestures in one learning setting, and could support studies of how these resources deploy simultaneously/in sequence and function as explanations, questions, signs of comprehension, requests for more information, etc.

(Un)intelligibility as a resource in deploying gestures in interaction

Michael Sean Smith

Using the video-recorded interactions of professional and novice geologists working in Yellowstone National Park, this paper examines the use of gesture in a complex multi- modal, multi-participant, multi-activity setting, and argues that the "intelligibility" of a gesture, i.e., the actor's sense or under-

standing of its meaning, is neither given nor transparent from the gesture alone. Rather, it is made accessible through the participants' successive actions as they pursue a joint multi-activity task.

The organization of such a process can be examined in one excerpt from the study showing an interaction between Darryl, one of the instructors, and Gina, a student. The sequence begins with Gina and Darryl breaking rock samples with a hammer on the ground. As Gina looks for a suitable place, Darryl suggests an alternative, points to the ground with his hiking pole and tells Gina "put it right here so it doesn't slip." Darryl's point initially seems unproblematic, for either him or Gina: she walks to the place Darryl pointed at and places the rock there. However, in response to what he sees her do, Darryl immediately bends down and repositions the rock, lifting, rotating it, and wedging it into a crevice beneath his prior point. His repair provides Gina an alternative way of seeing the rock-floor and the affordances it offers for their task. It also recasts his prior point and highlights in particular what he "intended" to show. While both participants seemingly act on the presupposition that the point is transparently intelligible, we instead see that its intelligibility is rendered visible through their subsequent interaction. Additionally, we see that through the repair Darryl can reveal for both Gina and himself that his point was previously (un)intelligible, or more accurately not yet fully intelligible to both participants' satisfaction. Lack of intelligibility then becomes a resource for Darryl in helping Gina shape her skills as a geologist.

As the analysis of this sequence and others shows, intelligibility, as a feature of the gesture, emerges continuously throughout the sequence. This would suggest that while actors collaborate in establishing and maintaining the intelligibility of their actions, the gesture itself, while potentially meaningful, does not guarantee its intelligibility. A gesture's "intelligibility" is then a historical and cumulative process, whereby its initial sense represents only one point along a course of action. As such, its development as a meaningful action is necessarily co-constituted by other features of the interaction, including the participants' relative understanding of their activity, community specific goals, and how tools and the environment can be utilized in pursuit of that activity.

Sensing technologies panel: Panel

Kinect gesture mapping with active skeleton

Patrick Gallagher, Zhuowen Tu, Baoyuan Wang, Baining Guo

Sensing Technologies Panel: Gesture and Action Recognition Abstract: The launch of Microsoft Xbox Kinect has made a big impact to the gaming industry; this sheds lights onto a wide variety of potential applications related to human computer interaction and action recognition. In this work, we propose an exemplar-based method to learn to map the initially estimated gesture to a new space, which can be used to drive an avitar in a new domain. We use the skeleton (medial axis) information to capture the main structure of an object, which has the particular advantage in modelling articulation and non-rigid deformation. Therefore, we then can use human gesture to drive to motion/action of a horse or a cat. We learn an inhomogeneous systematic bias by leveraging the exemplar information within specific human action domain. Our algorithm is illustrated on both joint-based skeleton mapping and tag prediction. Significant improvement is observed over the contemporary approaches.

Sensing technologies panel: Gesture and action recognition

James Hollan , Nadir Weibel

To inform the design of future interfaces for electronic medical records, designers must un- derstand the multimodal nature of the complex ecology of interaction between physicians, electronic medical record systems, and patients. However, characterizing multimodal activ- ity is difficult and expensive, typically requiring manual coding of hours of video data. As part of this panel, we will summarize our recent work on capturing and automatically coding body movement, speech, and gaze data from physician-patient interactions. In addition, we will discuss how a new generation of inexpensive commodity devices that integrate video cam- eras, depth sensors, and array microphones promise to enable not only automatic data capture but also novel interaction techniques involving whole-body interaction, gesture, and natural multi-touch interaction with information projected on desktops and other surfaces.

Immersion and interactivity

Todd Margolis

As interactive display surfaces continue to permeate society in public and private spaces, new approaches are required for individuals and groups to interact with these immersive display systems. If we consider immersion to be both a physical and social characterization of technology, we must investigate new forms of intuitive and responsive interaction modalities that can scale from personal media devices such as smartphones to room sized installations that surround groups within display surfaces. In this panel, I plan to discuss the multimodal design of a suite of interaction devices and platforms for effectively working and specifically collaborating with others using high resolution media display systems.

Analog interaction

Andy Wilson

While progress in sensing technologies and display hardware in the last ten years has been remarkable, there are still many challenges in designing so-called "natural" gesture-based interactions that are usable, useful and computable. The algorithms and representations we chose may have more to do with the success of future gesture-based interfaces than any new piece of hardware. I will discuss a number of issues that relate to the importance of representation, and how we might approach the problem of making our gesture-based interfaces more fluid and analog in nature. I will draw on our work at Microsoft Research developing prototypes that explore a variety of interaction techniques, form factors, algorithms and representations of gesture.

Space and viewpoint 1 : Talks

Fictive motion and gestures: Real discourse data from the TV news archive

Till Bergmann, Teenie Matlock

Fictive motion occurs when a static configuration is described in terms of motion, as in The road runs along the river and The mountain rises to 4,000 meters. In such cases, motion verbs are used, but objectively, nothing actually moves. Research on fictive motion suggests that sentences like these evoke mentally simulated motion, i.e., "movement" along a path (Matlock 2004a, 2004b; Matlock & Bergmann, in press), as well as modulate brain areas involved in motion processing (Saygin et al. 2010). Some work focused on fictive motion in mathematics has shown that using gesture to depict fictive motion facilitates mathematical thinking (N'o'oez 2008; Marghetis & N'o'oez 2013).

Despite a growing body of research on fictive motion, much of the work has focused on its linguistic parameters (e.g., Talmy 2000, Matsumoto 1996) or its comprehension in lab-based experiments. Ours is the first study to examine fictive motion in natural discourse, including gestures. The current work analyzes 98 instances of gestures that co-occur with fictive motion sentences, collected from the TV News Archive (http://archive.org/details/tv). Our presentation includes analysis of fictive motion sentences to describe a wide range of phenomena, including spatial scenes such as Interstate 15 runs from San Diego out to Las Vegas, where the subject is associated with motion, and The invasive species of plant that runs along the shore line, where the subject (species of plant) is not related to motion. Our analysis also includes cases with abstract entities as the subject, such as The interest rate hit the ceiling and went above it.

Our results suggest that speakers incorporate extra-linguistic knowledge in producing their fictive motion gestures. For example, we hypothesize that the direction of the gesture (sagittal or horizontal) depends on the perspective a person is taking while describing a spatial layout or scenario (more subjective versus less subjective). This is likely also to affect their mental simulation of the fictive motion scene, including whether they simulate themselves "moving" or an external object "moving". Some of the gestures discussed are cases in which the gesture provides information not expressed in concurrent speech, for example, a wavy gesture that accompanies The road goes to the mountain, where the manner of motion is specified just by the gesture.

Overall, our large-scale study shows that fictive motion is not just a linguistic construction or common form of non-literal language; rather, it is a general cognitive process that is expressed both in language and gesture. Studying how fictive motion gestures are articulated will help us reach a more coherent and comprehensive theory of fictive motion, and will thus contribute to the overall enterprise of understanding figurative language processing.

So the cat flies over where? Gesture/speech mismatch in a workplace storyboard design activity

Janet Blatter

Mismatches between what is gestured and what is spoken has been well documented in many tasks involving motion and frames of reference (FOR) (Alibali, Bassok, Solomon, Syc, and Goldin-Meadow, 1999); Melinger and Kita, 2004; Chu, and Kita, 2011; Kataoka, K., 2013). Examining the relationship between gesture and speech is particularly important in understanding discourse during authentic design activities (Suwa, and Tversky, 1997; Gero and Tang, 2001).

This presentation extends the research on gesture and mismatching occurring in complex, real-world, design discourse to a new domain, i.e., animation storyboarding. Storyboarding is the foundation of film planning, where key scenes of the intended film are sequenced on panels either on paper or digitally. A team of storyboard artists, the director and possibly the writer, reviews the board to assess if there is a viable direction for future development.

A viable storyboard depicts scenes that are structured and sequenced in a way that conveys a credible, coherent story, and a visually consistent, compelling film (Blatter, 2005, 2007). When collaborating artists review the boards, they must be able to "read" and identify problems in the board in terms of the story and how to present it as a film. Because each panel can show an angle different from adjacent ones, there are many spatial and logical inferences that the artists must make on behalf of the future film viewer. The function of the storyboard review is to make these inferences explicit.

The present research was part of a two-year, observational video-taped study with six (English-speaking) artists working on three productions. The presentation focuses on a problem in one 40-minute storyboard review. Utterance analysis (speech, hand/arm gesture, and drawing) of the videotape yielded 1800 sequential utterances of which 812 (and 243 concurrent speech/gesture) were devoted exclusively to solving one problem. The research question concerned the semantic relationship between gesture and speech, when are

they in semantic synch, and when do they diverge.

The findings revealed more than three times as many mismatches than agreement between gesture and speech when describing specific aspects of a motion event (i.e., location, direction, staging, etc.). Equally significant, almost twice as many mismatches than agreement occurred when the artists' referred to the story from a character's or observer's FOR within a 3D imagined space or the film from a viewer's FOR within a 2D plane).

We argue that mismatching points to mental simulation (Hegarty, 2004; Hostetter and Alibali, 2008) and gestural elaboration (Alibali, and Goldin-Meadow, 1993. Cassell, and Prevost, 1996; McNeill, 1995). Mismatching allows the artists to make salient aspects of the motion-event that are usually elided in English (Özyürek and Kita, 1999), and accommodates adopting multiple frames of reference (Tversky, 2011; Sassenberg and Van Der Meer, 2010).

This research has implications for extending current research in multimodal discourse to authentic, complex, workplace settings such as animation, filmmaking, and other time-based media design, where there are often conflicting goals and where mental simulation using gesture must account for specifying motion aspects in multiple spaces or FOR.

Gesture space revisited: From squares to spheres

Matthias A. Priesters, Irene Mittelberg

Gesture space has been shown to factor into the form and meaning of communicative postures and actions in various ways (e.g. Sweetser 2007). In this paper, we propose a revised theoretical and methodological approach to gesture space, here understood as encompassing a pragmatically conditioned hierarchy of spaces that flexibly constitute the communicative world inhabited and dynamically (co-)constructed by speakers (e.g. Sweeter & Sizemore 2008; Mittelberg 2010).

We depart from models that represent space as a set of squares or cubes, such as the center-periphery annotation schemes proposed by Pedelty (1987) and McNeill (1992). A speaker's gesture space, in our understanding, is the portion of space s/he actually utilizes for gesturing over a given discourse or conversation (Rodrigues 2010). It is seen as an evolving sub-space of the speaker's kinesphere (Laban 1966), that is, the part of real space (Liddell 2003) physically accessible to the speaker's hands. Segments of the kinesphere void of gestures are not considered to be part of a person's gesture space (Priesters 2012). McNeill (1992) showed that gesture spaces are characterized by their density structures. We understand these structures to be constituted by the traces of expressive gesture phases (strokes and holds; Kita et al. 1998), with individual speakers exhibiting highly idiosyncratic patterns in their gestural behavior (Priesters & Mittelberg 2013). As its name suggests, the kinesphere is spherical due to the anatomy of human limbs. Gesture (sub-)spaces also appear to be spherical in that movements of the hand normally revolve in arcs around the joints of the speaker's arms (e.g. shoulder, elbow, wrist).

Like other gestural form features, the spatial positions of gestures appear to be motivated (Priesters 2012; Mittelberg 2013b). Gestures are anchored in the speaker's body and kinesphere, produced with a certain degree of effort or force, but also coupled to the environment (Goodwin 2007) and oriented towards the interlocutor(s) and the ongoing interaction (Bavelas 1994). Furthermore, we draw on embodied conceptual structures such as image and force schemata (Cienki 2013; Mittelberg 2013a), some of which seem to prestructure both kinespheres and gesture spaces, e.g. the distribution of and relations between sub-spaces of higher semiotic density. In particular, spatial relation schemata such as LEFT/RIGHT, CENTER/PERIPHERY, FRONT/BACK, UP/DOWN, or PATH may underpin systematic uses of spatial segments to, for instance, reify discourse items or evoke relations between ideas, locations, or moments in time (e.g. Nüñez & Sweetser 2006).

Implementing these theoretical ideas, we introduce a novel technique engendering, based on both manual video annotation and automated analyses of motion-capture data, three-dimensional visualizations of gesture spaces. Results thus far encourage our view that gestures not only occur on a "flattened disk" (McNeill 1992), but on an ensemble of spheres surrounding the body and emanating from the joints of the upper limbs involved in communicative kinetic action (Kendon 2009).

Getting it right: Advanced Danish learners of Italian adopt speech and gesture l2 forms

Bjørn Wessel-Tolvig

Studies in motion events and gestures in Second Language Acquisition often show that, even though second language learners are able to produce correct L2 speech, they often apply typical L1 gestures thereby showing problems in reorganizing semantic representations and shifting attention towards different types of information (Choi & Lantolf, 2008; Kellerman & van Hoof, 2003; Stam, 2006). However, few studies show that language learners can to shift attention in both speech and gesture (Özyürek, 2002).

According to the typology proposed by Talmy (1985, 1991) Danish and Italian represent two different typological patterns. Danish, a satellite framed language, typically expresses path in verb particles and manner in main verbs. Italian, conversely, is categorized as a verb framed language where path is typically expressed in verb roots whereas manner tends to be encoded in subordinate clauses e.g. gerunds. However, typologies are not absolutely fixed. For example, Italian possesses manner verbs (e.g. rotolare-roll) observed to combine easily with verb particles (e.g. su-up, giúdown) (Folli, 2008; Iacobini & Masini, 2006). Nevertheless these constructions seem less frequent in Italian (Wessel-Tolvig, 2014).

This double strategy for lexicalization in Italian raises the question of how Danish learners of Italian come to encode information about motion in the L2. Are they more likely to select L1 typical (and grammatically correct) satellite framed forms or shift lexicaliztation pattern to more standard verb framed forms. And how will gestures reflect the choice of lexicaliztation.

Five Danish, five Italian and five advanced Danish learners of Italian narrated 8 motion scenes (Özyürek, Kita, & Allen, 2001; Wessel-Tolvig, 2013).

Consistent with findings for satellite framed languages, Danish L1 speakers predominantly express manner in verb roots and path in adverbial or prepositional particles in tight one-clause constructions and mainly produce one gesture (manner-path conflated gestures).

Although the Italian speakers predominantly express path in verb roots with manner, regularly, expressed in subordinate clauses, they also frequently use satellite framed constructions (verb + particle). With these one-clause constructions they mostly produce manner-path conflated gestures. When separating manner and path they tend to align gestures with semantically similar speech elements.

Surprisingly, the Danish L2 learners show remarkable consistency in expressing motion in "standard" verb framed forms. They express path in verb roots and either express manner in subordinate clauses, e.g. gerunds, or omit it. Their gestures reflect verb framed Italian patterns separating path and manner information, producing path gestures on path elements and manner/manner-path gestures on manner elements.

The results indicate, for the five L2 speakers, a reorganization of semantic representation and a shift in attention towards a uniform verb framed system which is opposite of their L1 system, but which does not really correspond to the reality of spoken Italian. Our results confirm the tendency observed in Özyürek (2002) with the modification that the form Danish learners seem to acquire is an abstract one. An explanation could be that verb framed constructions may be more frequent in Italian grammars and written texts (through which much university learning occurs) thereby limiting the knowledge and use of manner verbs.

Parallel Session 7 – Thursday, 10th 1:30 pm - 3:10 pm

Metaphors 1: Talks

The dynamics of multimodal metaphors in face-to-face communication.

Konrad Juszczyk, Ewa Jarmolowicz-Nowikow

When people express abstract concepts, such as emotions, attitudes and values they tend to use metaphors (Cameron 2010). According to Cameron, metaphor dynamics may result from the process of interaction, as one participant in a conversation responds to another. Recent research shows that metaphor not only helps in understanding and structuring experience (Lakoff and Johnson 1980; 1999) but also affects reasoning about time (Boroditsky and Ramscar 2002) and decision making (Thibodeau and Boroditsky 2011). Moreover, metaphor can be treated as multimodal phenomenon- abstract concepts are present not only in words, but also in gestures (Cienki 2008, Mueller 2008; Chui 2011) and graphics (Forceville 1996). Yet little is known about the way metaphor develops in conversation where all kinds of metaphor modalities appear. How one participant responding to the other enhances the development of gestural metaphors. How repeating of participant's words and mimicking participant's gestures influence the dynamics of multimodal metaphor in face-to-face communication.

To answer these questions we have recorded 96 coaching sessions and analyzed 80 of them. Coaching techniques and Clean Language questions (Sullivan and Rees 2008) were used in order to elicit multimodal expression of conceptual metaphor. Clean questions allow the participant to describe their personal experience with respect to their point of view and concepts used to construct their views (Tosey 2011). These techniques add to ecological validity of the presented study. In our experiments, the coach, while asking questions about career plans, repeates exact wording of the participant (hence parrot-phrasing) and parallels their gestures. Parallel gesturing in dialogue is often linked to imitation or mimicry and used to build or maintain rapport (Graziano, Kendon, Cristilli 2011). It also serves to facilitate the process by which participants sustain a common alignment to the conversational focus (Graziano, Kendon, Cristilli 2011). During the session, the coach also asks the participant to draw their metaphors of career, work and development.

Annotation of metaphorical gestures is based on MIP-G assumptions (Cienki 2012) plus structural and functional labels of NEUROGES coding system (Lausbserg & Slojtes 2009). According to MIP-G, the gestural metaphor is defined as a hand movement with stroke phase and abstract referent. In NEUROGES terms, metaphorical movements of hands are phasic or repetitive and performed to present form, motion quality and spatial relation (Lausberg 2013). In order to achieve high interrater agreement and reliability, the annotation was conducted by a group of ten trained annotators and supervised by two experienced researchers. Applying NEUROGES coding system and MIP-G assumptions to the identification of metaphorical gestures allows us for quantitative and qualitative analysis of metaphorical expressions.

The material analysed so far revealed variety of forms of parallel gesturing and multimodal metaphors expressions. It also suggests some level of coherence between verbal, gestural and graphical expressions of metaphors. In some observed cases, the number of participant's metaphorical expressions correlates with the number of coach's repetitions of participant's words and gestures. Further analysis will be conducted to confirm of current results by quantitative analysis on larger samples.

Where, what and how - Specifying gestures as a pathway to mathematics?

Christina M. Krause, Angelika Bikner-Ahsbahs

While in many disciplines students deal with objects that are directly accessible, the case is crucially different for mathematics: Mathematical objects are not tangible. Although their handling can only be mediated by representations (Duval 2000, 61) they must not be confused with them (Duval 2006, 107): Mathematical objects are abstract individual constructions. In math classes mathematical objects often are established within social interactions, even without the teacher. This raises the question how a shared understanding of such abstract concepts among students is possible. A shared understanding may be accomplished by establishing "conceptual pacts" (Brennan & Clark 1996), that is, shared conceptualizations in discursive situations. Research has already shown that gestures take major part in this (Arzarello, Paola, Robutti & Sabena 2009, Edwards 2005, Radford 2003, Goldin-Meadow 2003, 2010, Goldin-Meadow, Cook & Mitchell 2009, Cook 2011) but little is known about how this happens.

In the current study, we research the issue of how gestures contribute to building mathematical conceptual pacts on a highly abstract level.

This has been investigated by an empirical study by means of teaching experiments, which induced dense and highly abstract epistemic processes of building mathematical concepts. For this, three couples of high achieving students of grade 10 solved three problems from diverse and abstract mathematical areas. The epistemic processes of solving the problems, each lasting from 60 to 160 minutes, have been videotaped considering three perspectives in split-screen. All data has

been transcribed focusing on verbal and non-verbal actions. The students' mathematical discourse was analyzed using the semiotic sequence analysis (Bikner-Ahsbahs 2006, 161f.) in a fine-grained way, linking speech, representations and the use of gestures. An epistemic action model (Bikner-Ahsbahs & Halverscheid 2014) allowed identifying the crucial moments of building new mathematical entities by the epistemic action of structure seeing, the nucleus of conceptualization. The core units of gesture analysis concerned strokes and post-stroke holds (McNeill 1992, p.25, Kendon 2004, 112) and their relations to other semiotic resources in the current epistemic process.

Results show that establishing shared conceptualization benefits from specification gestures, i.e. gestures that provide information additional to that conveyed in speech. This is done by gestures specifying aspects of mathematical objects within a certain representation. Our analyses reveal four different kinds of specifying gestures: specifying aspects of mathematical objects through location, sort and style in the representations at hand, and specifying relations between mathematical entities as displayed. These specifying gestures and the mathematical concepts under construction can be shared on three referential levels, the one of inscription, the one of the interaction space with indexical gestural reference to the inscription, and finally the level of the mathematical gesture space in which mathematical meaning is blended from inscription (Yoon, Thomas & Dreyfus 2011).

Left-hand gestures enhance metaphor explanation

Paraskevi Argyriou, Sotaro Kita

Research suggests that gestures influence cognitive processes, but the exact mechanism is not clear. Additionally, it has been shown that when a linguistic task (metaphor explanation) involves the right brain hemisphere, the left hand becomes more gesturally active. We hypothesized that gestures with a particular hand activate cognitive processes in the contra-lateral hemisphere. We examined whether gestures with the left hand enhance metaphoricity in verbal responses. Results showed participants produced more metaphoric explanations when instructed to produce gestures with their left hand as compared to the right hand or not gesture at all. In addition, we measured the mouth asymmetry during metaphorical speech to determine individual differences in right-hemisphere involvement in metaphor processing. The left-side mouth dominance, indicating stronger right-hemisphere involvement, positively correlated with the left-hand-over-right-hand advantage in gestural facilitation of metaphorical speech. We concluded that left-hand gestures enhance metaphorical thinking in the right hemisphere.

Tiny numbers are actually tiny: Precision grips map onto implied quantity

Bodo Winter, Marcus Perlman, Teenie Matlock

We frequently talk about numbers in terms of size, for instance, "this is a large number" or "this is a small number." In Conceptual Metaphor Theory, such language about SIZE is thought to reflect the speaker's underlying conceptualization of NUMBER. Here, we support this idea by showing that when people talk about "tiny numbers" in spontaneous discourse, they frequently use precision grips, as if metaphorically seizing a small object. Using the TV News Archive (https://archive.org/details/tv), we analyzed the complete set of occurrences of the phrase "tiny number" in the corpus (which spans three years of American Television). Within this set (N=156), there were 63 visible gestures from 61 different speakers. More than half of these gestures (N=35) involved a precision grip, far more than other hand shapes.

We discuss our data in light of Kendon's (2004) distinction between ring-type gestures (index finger and thumb together) and grappolo-type gestures (all fingers together). Of the 35 precision grip gestures we analyzed, 26 were ring-type gestures, but only 6 were grappolo. Besides the discursive functions outlined by Kendon, our data suggests that R-type gestures map SMALLNESS onto quantity. G-type gestures are perhaps dispreferred because the corresponding movement could be used for grasping a somewhat larger quantity of objects. The association between size and quantity is further highlighted by instances in which speakers express quantity differentials (see Fig. 1), and use opening hand and arm movements to talk about large quantities.

We also discuss our results in connection to work on metaphorical gestures (Cienki & Müller, 2008) and on numerical cognition (Lindemann et al., 2007). We argue that the consistency with which speakers use precision grips to indicate quantity is one key source of evidence that shows that we use space to think about numbers, and, more generally, that mathematics is grounded in everyday embodied experience (Lakoff & Núñez, 2000).

Development 2 : Talks

The role of iconicity in linking language to the world

Jenny Lu, Pamela Perniss, Gary Morgan, Gabriella Vigliocco

Learning language requires linking linguistic form to referents in the world. A central problem is explaining how children manage to learn labels that are linked only arbitrarily to referents and how they learn to make correct associations despite the ubiquitous presence of multiple referents. Solutions to this two-fold problem of referential ambiguity include the infant's statistical learning abilities, visual isolation of referents, inferring communicative intentions, and caregivers' use of pointing gestures to establish attention to a referent while producing its label. Here, we provide initial evidence for the use of iconicity resemblance between form and meaning as an additional mechanism for reducing referential ambiguity.

Specifically, we hypothesized that caregivers might support referential mapping by modifying language in ways that make the link between form and meaning more transparent, i.e. by highlighting action/perception-based properties of referents. In addition, we hypothesized that iconicity could play an important role in identifying a referent from language even when the referent cannot be directly attended to. In such cases, e.g. when parents talk about things that are not in the here-and-now, the use of iconicity to evoke action/perception-based properties of referents may be particularly useful in helping to bridge form and meaning.

Previous research has argued against a significant role for iconicity in supporting the referential mapping process. However, the presence of iconicity in children's early language input has not been investigated in detail. Moreover, there has been no investigation into the way different strategies (e.g. iconicity and deixis) may work in tandem to scaffold referential mapping across different contexts of communication, such as whether objects being talked about are present or not in the immediate environment. In this study, we investigate caregivers' use of iconicity in creating word-to-world mappings in British Sign Language (BSL). In addition, we ask whether the communicative context (i.e. presence vs. absence of referents) influences how caregivers use language to establish reference, and compare the use of iconicity vs. deixis.

We start with sign language, because of the high potential for iconic mapping afforded by the visual modality. In a within-subjects design, deaf adult signers (N=10) were given four novel toy sets and asked to imagine playing with and talking about each set with their child in two conditions (toys present; toys absent). We found that parents used iconic modification of both manual and non-manual components of signs (i.e. emphasizing perception/action-based features of referents) to a greater extent when objects were absent than when they were present, where instead deixis was relied on more. In contrast, non-iconic modification (e.g. a larger sign produced for purposes of emphasis or attention-getting, but that does not highlight perception/action-based features of a referent) was used equally in both contexts. Our results suggest that caregivers use iconicity to support the referential mapping process, and that its use is modulated by communicative context (i.e. presence vs. absence of referents). Iconicity may thus be a powerful tool for reducing referential ambiguity and scaffolding child language acquisition.

Code-blending in bimodal bilingual development

Ronice Quadros, Deborah Chen Pichler, Diane Lillo-Martin

Children who are exposed to a spoken language and a signed language can become bimodal bilinguals. Like adult bimodal bilinguals (Emmorey et al. 2008), children produce a variety of structures reflecting one or the other language, and most interestingly, structures reflecting the influence of both languages. The latter include cases of cross-linguistic influence (code-mixing), code-switching, and code-blending. Code-blending is a unique reflex of the bimodal bilingual's option to produce (portions of) a linguistic message using both modalities simultaneously.

In this presentation, we focus on instances of code-blending in the spontaneous production of bimodal bilinguals (hearing children with Deaf parents) from two language pairs: English + American Sign Language (ASL), and Brazilian Portuguese + Brazilian Sign Language (Libras). We will report on data from children ages 1;04-3;09, and their adult interlocutors.

Our model considers code-blending to be one possible outcome from a derivation that freely makes use of linguistic elements from both languages. The derivation is constrained by the need for selected elements to be appropriately licensed. Each utterance produced, whether unimodal or bimodal, reflects the derivation of one proposition.

Looking more deeply at the quality of bimodal productions, we find several types. In speech target sessions, the majority consist of speech+point V an option also available to monolinguals in a spoken language environment. In both speech and sign target sessions, blending types include fully and partially bimodal utterances which are almost always congruent.

Our quantitative analysis reveals that the majority of productions are unimodal, with greater bimodality in sign language target sessions. There is variability in the amount of bimodality, but this is not clearly a developmental issue, as the proportion does not change steadily with age.

To explore our hypothesis that blending utterances reflect the output of a single computation, we analyzed the amount of overlap between the speech and the sign. The cases of most interest are those in which there is a (partial) mismatch. We find three types: i) cases of apparent mismatch in the number of utterances (e.g., one sign corresponding to three spoken utterances); ii) timing mismatches (speech and sign are not produced at the same timing beat); iii) non-redundancy (part of the message is conveyed in each language, but neither contains the full message).

All cases of mismatch were further analyzed. The type (i) mismatches involve holding or repeating a sign or word, and are used as a conversational strategy for holding attention, maintaining the topic, or repairs. The type (ii) mismatches involve immature coordination between manual and vocal outputs, with repetition used to repair ill-coordinated timing. All of the type (iii) mismatches are cases where a full single proposition makes use of both spoken and signed pieces; none reflect simultaneously produced independent propositions.

We conclude that children are different from adults in that they are still developing coordination, but otherwise they make full use of the possibilities made available in bimodal bilingualism. In particular, they may combine aspects of both languages as the output of a single computation.

A second language learner's thinking for speaking in her l1 and l2 after fourteen years: Verb framed, satellite-framed or in between?

Gale Stam

Slobin (1991) has proposed that children learn a particular pattern of thinking for speaking in first language (L1) acquisition, and Stam (1998) has argued that second language (L2) acquisition often entails learning a different pattern of thinking for speaking. Cross-linguistic motion event research has shown that Spanish speakers and English speakers have different patterns of thinking for speaking about motion linguistically and gesturally (for a review, see Stam, 2010). Spanish speakers express path linguistically on verbs, and their path gestures tend to occur with path verbs, whereas English speakers express path linguistically on satellites, and their path gestures tend to occur with satellite units.

Stam (2006) has shown that the English narrations of Spanish learners of English have aspects of their first language (Spanish) and aspects of their second language (English) thinking for speaking patterns in both speech and gesture. She has further shown that these patterns can change (Stam, 2010): an L2 learner's thinking for speaking about path in English became more native-like, but her thinking for speaking about manner did not. This raises several questions: Do learners' L2 thinking for speaking patterns continue to change with regular use of the L2. Does this also affect their L1 thinking for speaking patterns.

To investigate these questions, motion event narration data gathered in 2011 was compared with data from 1997 and 2006 to examine how the learner's expression of path and manner linguistically and gesturally in her L1 (Spanish) and L2 (English) changed. The results indicate that the learner's gestural expression of path continued to change in both her L1 and L2 and that her gestural expression of manner changed in her L2 between 2006 and 2011. This change suggests that manner is not resistant to change after all (Slobin 1996; Stam 2010) and thinking for speaking is not static.

Age related changes in the processing of bimodal production between 3 and 11 years

Jean-Marc Colletta, Catherine Pellenq, Ali Hadian Cefidekhanie

The research devoted to language acquisition from a multimodal perspective recently brought to light several new and important findings on the way language abilities are tightly linked to gesture production, at early stages as well as at later stages. Focussing on early stages of the child's development, studies mostly focussed on the semiotic content of speech, the interrelation between gesture and speech and the way gesture and word combines predict some aspects of early language acquisition (Fasolo & DÓdorico, 2012; Guidetti, 2002; Valloton, 2010). Focussing on later stages, studies mainly investigated the way gesture and speech are synchronized and related on both a semiotic and a functional ground, and described how gesture production evolves and contributes to narrative among other types of discourse performance (Graziano, 2009; Kunene, 2010; Sekine & Furuyama, 2010). However, how the child actually processes speech and gesture information in language production remains a largely unstudied domain.

Our study examines age-related changes in oral narrative discourse of 140 French school-children aged 3-11 years who elaborate a story based on a short animated film. All children's language abilities were assessed. Language and gesture data were transcribed and coded using ELAN as an annotation tool. Basing our methods on the literature devoted to the study of speech rate and the planning of speech (Ryan, 2000; Pavao Martins et al., 2007), we measured all relevant variables (syllables, words, clauses, gesture strokes) in order to calculate speech rate as well as gesture rate, and to track changes in the informational content of prosodic units (speech segments marked by pauses and an intonation contour) and syntactic units (clauses). We find an increase of speech rate as children grow older, as well as a significant increase of gesture rate. Our results also show that age has an effect on the density of informational content both of prosodic (speech segments) and syntactic units (clauses). As children grow older, they produce longer speech segments and they load clauses with more verbal and gesture content.

These results are interpreted in terms of the age-related changes in underlying cognitive abilities they reveal (representation, syntactic parsing, short term memory and speech planning). They are discussed in line with the available models of language production in adults (de Ruiter, 2000; Kita & .zy.rek, 2003; McNeill, 2005) and the above-mentioned developmental studies.

Gaze in gesture and signing: Panel

Maintaining multiple viewpoints with gaze

Eve Sweetser, Kashmiri Stec

Gaze is a powerful and basic marker of attention- and director of joint attention, as small children very early learn (e.g. Tomasello 1999). If two people are talking, and one suddenly looks aside at something behind her addressee, the addressee is likely to turn to see what object behind him is of interest to the speaker. But when multiple mental spaces are involved- for example, when a speaker is describing a past conversational interaction- every aspect of the speaker's gesture, including gaze, is potentially ambiguous, or more than ambiguous. Should we attribute the relevant gesture to the speaker herself. Or to one or another of the depicted conversational participants. How can we tell.

Gaze is like other aspects of gesture in that speakers regularly use it both in their own behaviors and in embedded depicted spaces. Quite frequently, while immersed in the depiction of narrative content, the real speaker's gaze simply indicates the gaze of a participant in that space- and may alternate between depicting one participant's gaze and another's. This is rather similar to the ways in which hand behaviors and body postures can be understood as depicting aspects of described participants' bodily behaviors, whether sequentially or simultaneously (e.g. Dudis 2004).

However, gaze may be different from body posture and hand location in that it is easier to use to mark transitions between mental spaces (Fauconnier 1997; Fauconnier & Turner 2002). In both signed language and co-speech gesture, there is the possibility of moving gaze (and often facial orientation) back towards the real-world interlocutor, while leaving hands in the physical space allocated to the embedded content space (and possibly leaving the trunk directed towards that space) as a "buoy" maintaining that embedded space's presence (Stec & Sweetser 2013). Thus two bodily viewpoints are maintained simultaneously, one by the trunk/hands and one by the head- and most often, in this case, the higher viewpoint is the one correlated with gaze.

We want to suggest that one reason why gaze is so good at this is the speed and effort of shifting viewpoint- it is extremely fast-moving, with little effort. In fact, it is a meta-level pragmatic marker of "checking" with another participant, for this very reason. In this paper, we will document the ways in which gaze is used to signal multiple perspectives, and how it coordinates with co-speech gesture to do so.

Viewpoint in signed discourse: The privileged status of the signer's body and gaze

Elisabeth Engberg-Pedersen

McNeill (1992) distinguished two types of viewpoint in gesture based on the involvement of the speaker's body: character viewpoint, where the speaker's body is incorporated into the representation, and observer viewpoint, where the speaker's body is not incorporated. Based on this distinction, Parill studied " hand and body gestures that simultaneously express multiple perspectives on an event or scene" (2009: 271), so-called dual viewpoint gestures (McNeill 1992, Parill 2009). Parill explicitly excluded facial expression and gaze phenomena.

Signers may represent several referents simultaneously by means of their two hands, their body and their facial expression (Engberg-Pedersen 1995, Liddell's (2003) notion real space blends and Dudis' (2004) notion body partitioning). When two articulators represent two distinct referents simultaneously, one articulator may be active and represent foregrounded information in relation to the information represented by the other articulator (Engberg-Pedersen 2011). Two major contrasting constructions involving the signer's body and hands have been identified, one where the body and the hand(s) are aligned and represent the agent of the action, and one where one or both hand represent(s) the agent and the body represents the patient. The latter has been compared to passive constructions in spoken languages, i.e., in these constructions the referent represented by the signer's body is akin to the referent of a patient subject in spoken languages (Janzen, OśDea & Shaffer 2001, Janzen 2004). The constructions where two different referents are represented simultaneously by the hands and the body, would undoubtedly be described as dual viewpoint gestures (or signs) in Parill's terminology. However, in signing the major distinction seems to be constructions where the signer's body represents the agent (corresponding to active clauses) and constructions where it represents the patient (corresponding to passive clauses). Furthermore, we need to include the signer's gaze direction to get the full picture.

Simultaneously with and to some extent independently of the other articulators, the signer's gaze may be used in two major ways, either to monitor the communicative act, in which case the gaze direction should be ascribed to the signer/narrator, or to imitate a character's gaze direction (Engberg-Pedersen 2003). When the gaze is used to monitor the communicative act, it is used to stay attuned with the receiver and maintain the signer's conversational turn (eye contact and brief interruptions in eye contact with looks in no specific direction just before or at major syntactic boundaries), to track referents (brief glances in the direction of a locus representing a referent, usually at the beginning of a new sentence), and to attract the receiver's attention to a particularly complicated configuration designated by the hands (more prolonged looking at what the hands are doing). When signers imitate a character's gaze direction, the viewpoint is clearly with this character.

In my talk, I shall give examples of the different combinations in adult and child signed narratives and demonstrate that the proper use of gaze is the last aspect to be acquired by children.

Gestural links to grammatical eyegaze during production of American Sign Language verbs

Robin Thompson, Karen Emmorey, Clifton Langdon

We investigated whether English-speaking non-signers are sensitive to the directed eyegaze that accompanies verbs in ASL. Discourse referents in ASL are regularly associated with spatial locations, and signers direct manual verbs toward these locations to indicate referents. An "agreeing" verb directed to spatial locations specifies the grammatical role of the referents, i.e., subject and object. Signers additionally direct their eyegaze toward the location associated with a referent, marking it as syntactic object for both "regular" agreeing verbs (subject marked manually first, object second) and "backwards" agreeing verbs (object first, subject second) (Thompson, Emmorey, Kluender, 2006). Not all verbs mark agreement.Plain verbs are not marked manually or with gaze,

but are sometimes displaced in space.

While rules constraining the placement of hands and eyegaze in ASL are complex, there are similarities between using spatial locations to mark agreement and using co-speech gestures including eyegaze. When speaking, concepts can be visualized in space and gestures are sometimes directed toward spatial locations to indicate referents (McNeill, 1992). Speakers also gaze toward these locations when referents become the focal point of the discourse with gaze similar to a manual gesture (Streeck, 1993; Gullberg and Holmqvist, 1999).

We investigated potential parallels in gaze accompanying gesture and sign by asking whether non-signers (along with ASL signers) are sensitive to eyegaze occurring with agreeing verbs. Participants were asked to match a picture of a signer's eyegaze to one of two pictures of an ASL verb.Four different verb pairings were used: 1) two agreeing verbs one moving right, one moving left; 2) an agreeing and a backwards verb; 3) an agreeing and a plain verb; 4) an agreeing verb and a spatially-displaced plain verb. The findings indicate non-signers are sensitive to gaze direction in most cases. Gaze choice for native ASL signers (Deaf=22, hearing=16) and English speaking non-signers (N=21) was more accurate when correct gaze choice matched hand movement (conditions 1, 3). While both groups were less accurate for condition 4 (compared to 1&3), there was still a strong preference for the agreeing verb over the spatially-placed plain verb for both groups, suggesting choices were driven by perceived directionality of the sign/hand, not the end location. When grammatical eyegaze could not be predicted based on direction of movement (condition 2), non-signers were significantly less accurate selecting gaze direction for backwards agreeing verbs compared to regular agreeing verbs, while signers performed equally well on both conditions.

The results provide a broader understanding of language communication systems (both signed and spoken) and the degree to which spatial information (from hands and eyes) is integral to comprehension regardless of spoken/signed modality. The possible influence of gesture on the emergence of the grammatical system for ASL is intriguing. It is possible that grammatical eyegaze in ASL has its roots in the gaze that co-occurs with manual gestures. However, only signing participants showed sensitivity to eyegaze marking the syntactic object for backwards verbs when gaze did not match the hand movement. Thus, knowledge of the linguistic system is ultimately needed.

Interaction 4: Talks

Manipulation of objects in proposal sequences as a resource to fine-tune the interactional outcomes

Chiara M. Monzoni, Melisa Stevanovic

Previous research has demonstrated how participants' bodily behaviour constitutes a key part of the multi-layered action-formation machinery (Levinson, 2013; Streeck, 2009; Mondada 2009; among others): this is even more true when it comes to the deployment of actions and responses which can be carried out through embodied actions only, such as responding to (proximal) requests (cf. Rossi, 2012).

Basing our analysis on Finnish and Italian data, and using a conversation analytic approach, we will consider how participants use embodied behaviours more specifically, those involving the manipulation of objects V as resources during proximal proposals sequences. In proposals, speakers prototypically name courses of action suggesting these to be realized by their recipients, while presenting that realization as contingent on the recipients' acceptance; thus, establishing a relational symmetry between the proposer and the recipient (unlike more imposing actions, such as commands). Here, embodied actions are key to the compliance (or non-compliance) of proximal proposals, since proposed actions are to be carried out immediately.

Given the very nature of proximal proposals sequences, we will demonstrate that the development of embodied activities through the manipulation of objects is more relevant than the development of verbal ones. The manipulation of objects by proposers and recipients sets up an apparent discrepancy between the courses of embodied and verbal actions, which has consequences for participants' larger interactional project(s). Recipients may accept proposals verbally, while avoiding complying immediately with the proposed course of action. Instead, they may engage in other concurrent activities involving the manipulation of (other) objects, thereby postponing the (verbally) agreed course of action V if not ignoring or rejecting the original proposal altogether. Proposers counter this type of recipients' behaviour by manipulating those very objects that are key for the actualization of the proposed courses of action. In this way, the original proposals are kept in play (i.e. proposers show that their original proposals are still relevant for the interaction), while the manipulation of objects can help them to construct further opportunities for themselves verbally to re-issue their original proposals. Hence, the hiatus created between the embodied medium and the conversational one is not accidental, but constitutes a resource employed by participants to fine-tune the courses of the unfolding of interaction and the final outcome of proposals.

Compared to more explicit activities for instance, to those in which resistance to a proposal can be expressed verbally or a proposal can be turned into a command participants' use of the above-mentioned embodied actions might at first appear as something subtle and mitigated. Even though this might be true for some cases, in other instances, embodied actions seem to be quite forceful and imposing resources to push for the actualization or non-actualization of some courses of action. This observation will be discussed in relation to preference organization.

Gesture, gaze, and framing in situated activity

Simon Harrison, Robert F. Williams

For Goffman (1974), frames are the principles of organisation which govern events K and our subjective involvement in them (pp. 10-11).In this paper, we analyze the role of the hands and eyes in structuring, shifting, and blending frames in five minutes of situated activity among lifeguards on a beach in southwestern France.

Previous frame analyses of discourse have shown how speakers use language to structure frames, shift between frames, and blend frames together as they interweave different activities in context (Gordon, 2008; Tannen, 2006).Our analysis extends this work by examining the roles of gesture and gaze-shifts in these processes.

In our data, one lifeguard attempts to teach two co-workers how to find south without a compass while all three simultaneously monitor the swim zone.Here a social frame for instruction is interleaved with a work frame for lifeguarding.Attention is divided between the ocean and four spaces in which the primary speaker represents aspects of the activity being taught: the sand, the surroundings, his wristwatch, and the air in front of his body (Authors, 2012).Our paper elucidates the shifts in eye gaze and gesture that mark transitions between the two task frames and among the various representational spaces whose contents are being interrelated.

To analyze these characteristics of the discourse, we used ELAN to code the data for task frame, spoken language, eye gaze, and various aspects of gesture, including space, mode of representation, and environmental coupling (Authors, 2012, 2013; Goodwin, 2007; M.ller, 1998; Streeck, 2009).We then examined how gaze and gesture relate to framing. Our findings provide evidence that shifts in eye-gaze accompany shifts in frame, that changes in mode of representation reframe the discourse, and that speech-gesture combinations blend frames together.

Parallel Session 8 – Thursday, 10th 3:30 pm - 5:10 pm

Development 3 : Talks

Intercultural variation in the use of representational gestures in expressing motion events by Zulu and Sothu Speaking Children

Ramona Kunene

To study age related changes in the way children and adults gesture while narrating leads us to better estimate the relative weight of social and cognitive factors in narrative development (Berman, 2004). Studies on late language acquisition have shown that from 9 years of age and onwards, narratives gain in linguistic complexity and children increase their frequency of co-speech gesture use to represent the narrated events and characters, to maintain the internal coherence of the narrative, and to mark the transitions between the account of events and the commentaries (Colletta et al., 2010; Kunene, 2010).

Cross-linguistic studies on the influence of motion events in oral speech and co- speech gesture have steadily shown how languages constrain lexical and syntactic resources in different languages (Talmy, 1985; Kita, 2009; Kita & Özyörek, 2003). Literature also shows that children are sensitive to the typological constraints applied to their languages and express motion events accordingly.

There is currently little literature on the multimodal expression of motion events in Bantu languages and even less on how Bantu language children express motion events. This paper seeks to compare the variation in the expression of spatial information of Sotho and Zulu, which are Bantu languages spoken in South Africa.

In an empirical language production task, data of Sotho and Zulu oral narratives was gathered from 3 children groups aged 5 years, 9 years and 12 years corresponding to early, middle and final primary school levels. Oral narratives were also collected from adult speakers of the two languages in order to measure the developmental trajectory. In a classic method for collecting audio and visual data (McNeill, 1992), children watched a brief animated cartoon and were asked to recount what they had seen to an interviewer. Narratives were filmed and coded on ELAN software for both speech and gesture.

Results reveal that both Zulu and Sotho are verb-framed languages according to Talmy's (1985) classification and this is also expressed through co-speech gesture. There was a high frequency of representational gestures expressing manner rather than path, similar to the cross-linguistic study of Japanese and Turkish compared to English by 'ozy'orek et al, (2008). The results also revealed the frequency of representational gestures increases with age in both languages. This study brings further evidence to the findings of motion events being expressed parallel in both speech and gesture and also brings some insights on lesser-studied languages such as Sotho and Zulu.

Encouraging right handed pointing improves children's performance in a linguistic task

Katherine Mumford, Sotaro Kita

Research has shown a strong association between language development and right handed gesturing (e.g. Cochet, Jover & Vouclair, 2011). For example, infants with a larger vocabulary have been shown to produce more right handed pointing gestures than infants with a smaller vocabulary (Mumford & Kita, 2013, Esseily, Jacquet & Fagard, 2011). Further, in childhood, gesturing has been shown to help several linguistic processes. For example, gesturing can aide word retrieval when in a tip-of-the-tongue state (Pine, Bird & Kirk, 2007) and seeing other people gesture has been shown to aide word learning (Mumford & Kita, in press; McGregor, Rohlfing, Bean & Marschner, 2009).

The current study aimed to combine these two bodies of research by investigated whether right hand gesturing may facilitate word learning more than left handed gesturing. Such a right hand advantage is expected due to the overlap of neural substrates for speech and gesture processing in the left hemisphere (e.g. Skipper, Goldin-Meadow, Nusbaum & Small, 2007; Kimura; 1973).

Specifically, the current study investigated whether manipulating the hand with which children were allowed to use in pointing tasks, would affect their performance (in either a verb learning task (linguistic) or a memory task (non-linguistic)). One-hundred-and-sixty 3-year-old children were tested, Twenty-nine were excluded including eight who showed a strong right hand bias for manual activities (writing and opening a screw lid). Therefore, only the data from right handed children who were assumed to be developing language dominance in the left hemisphere was analysed.

The results showed that while the hand manipulation had no effect on performance in the memory task, using the right hand significantly increased performance in the verb learning task compared to using the left hand. The results are the first to suggest that manipulating hand choice for gesturing may be able to influence the performance of a linguistic task. Specifically, right handed gestured may have an especially close relationship to language development due to the overlap in the neural networks.

It should be emphasised that this study only included children who were naturally right handed and it does not, in any way, support the practice of encouraging left handed children

to use their right hand. It would be interesting for future research to look more closely at language lateralisation and the benefit of using the ipsilateral hand for gesturing in linguistic tasks.

The role of visual modality in development of encoding spatial relations

Beyza Sumer, Pamela Perniss, Inge Zwitserlood, Asli Özyürek

Acquisition of viewpoint-dependent spatial relations (e.g., left, right) is reported to be a late aspect of language development. However, previous studies have focused on speech only, thus are not informative about the development of these expressions in the visual-spatial modality, such as in co-speech gestures and sign languages. Sign languages and gestures are interesting in this domain due to the modality's visual-iconic and embodied affordances: spatial relations can be expressed in an analogue manner in the space in front of the body or by using body-anchored signs (e.g., tapping the right and left hand/arm to mean right and left). Previous research has reported that children learning sign language lag behind speaking children in comprehending viewpoint-dependent spatial relations. However, production studies comparing signing and speaking children are lacking. Furthermore, past studies on speaking children have failed to systematically consider the visual-spatial modality (i.e., co-speech gestures) in their analyses.

Here we attempt to track the developmental pattern of spatial language used to encode viewpoint-dependent relations in Turkish Sign Language (TID) and Turkish. Deaf children in two age groups (preschool-age: 4-6 years & school-age: 7-9 years; N=10 in each group) acquiring TID natively and age-matched hearing children acquiring Turkish (N=10 in each age group) described pictures depicting two objects localized with respect to each other on the left/right axis (e.g., pen to left of paper). Their descriptions were compared to adult descriptions in each language.

TID-signing adults mainly employed classifier predicates providing analogue representations of the scenes, but they also used body-anchored lexical signs (i.e., left, right). Deaf children were adult-like in frequency of use of both classifier predicates and body- anchored lexical signs. Turkish-speaking adults either used a general relational term (e.g., Kalem kag'od'on yan'onda"pen is at the side of the pape") or employed a viewpoint-dependent spatial noun (e.g., Kalem kag'od'on solunda "pen is to the left of the paper") to encode the spatial relations. Turkish-speaking children of both age groups used the general relational term as frequently as adults, but they used the spatial nouns specifying viewpoint significantly less than the adults (and preschool-age children never used them). Additionally, 7-9-year-old children used co-speech gestures to provide visual representations of the spatial relation between entities more than adults and preschool-age children, especially when their speech included only the general relational terms.

These results imply that learning to express viewpoint-dependent spatial relations might be facilitated through use of the visual-spatial modality. In the case of sign language, deaf children learn the adult patterns V i.e., the analogue spatial representation using classifier predicates and body-anchored lexical signs encoding viewpoint V earlier than their speaking peers who need to learn spatial nouns that do not exhibit such iconic, body-anchored form- meaning correspondences. This advantage seems to manifest itself in production rather than in comprehension. Additionally, the observed increase in the use of co-speech gestures by 7-9 year-old Turkish-speaking children suggests that the visual-spatial modality might pave the way to learning spatial language that requires encoding viewpoint-dependent relations in speech.

Argument structure in a newly documented village sign language

Rabia Ergin, Naomi Berlove, Deniz Ilkbasaran, Irit Meir, Ariel Cohen-Goldberg, Ray Jackendoff

Central Taurus Sign Language (CTSL) is a village sign language that emerged as a result of recessive deafness in two villages in the Central Taurus Mountain region south-central Turkey. We have identified roughly three generations of signers. Approximately 15 deaf people from the second and third generations currently live in one village and 13 in the other (3-4% and 0.5-0.6% of the population, respectively). Due to cultural, geographical and financial circumstances, only five deaf villagers received some formal education and learned Turkish Sign Language (TID). A native TID signer verified that despite minimal influence from TID the two are quite distinct.

Like all village sign languages, CTSL provides a window into the human capacity for language and can help answer the question: What is it that humans are capable of creating from scratch in absence of a linguistic tradition'o As one of a growing number of village sign languages that have been identified, CTSL represents an opportunity to draw comparisons among village sign languages and to begin to make modest generalizations about language emergence. Preliminary analyses reveal some similarities between CTSL and other village sign languages. Like Al-Sayyid Bedouin Sign Language (ABSL), signers point to absolute locations (Aronoff et al., 2008), and make frequent use of compounds (Meir et al., 2010) and suprasegmentals like eyebrows which accompany the manual signs (Sandler et al., 2011). Surprisingly, CTSL signers sign with their feet in addition to their hands (e.g., for naming colors).

We present an analysis of the strategies used in CTSL to disambiguate arguments in a given proposition when there are two animate semantic characters (e.g., The man threw the ball to the girl). Signers described a set of short video clips (Sandler et al., 2005) to an interlocutor, who then selected the matching picture from an array of three pictures.

Signers often first attempted to use word order without morphological marking (e.g., MAN GIRL BALL THROW), but this was rarely understood by the interlocutor. Like ABSL (Meir, 2010; Padden et al., 2010), ISL , (Meir, 2010) and NSL (Senghas et al., 1997), CTSL signers sometimes disambiguated semantic roles using pragmatics and divided the proposition into two sentences (e.g., MAN THROW / GIRL CATCH.). They also used two additional syntactic strategies. Sometimes they would use spatial inflection by setting up a location to refer to each of the semantic characters, and then use verb agreement to mark agent and recipient/patient syntactically (e.g., MAN LOC1, GIRL LOC2, LOC1-THROW-LOC2), where the path of movement matches the direction of transfer between the agent and recipient. More frequently, they would identify themselves as one of the characters in the discourse, and again use verb agreement to mark themselves as the agent or the recipient/patient of the event (e.g., ME MAN, LOC1 GIRL, ME-THROW-LOC1). These findings suggest that CTSL signers use both syntactic and semantic strategies for encoding argument structure when two semantic characters are animate.

Robotics: Talks

Gestural alignment between real and virtual humans

Kirsten Bergmann, Stefan Kopp

There is a large body of work demonstrating inter-personal sensitivities in both verbal and nonverbal behavior in natural social interaction, leading often to coordination and alignment between interlocutors (cf. Kopp 2010). With respect to gestures, however, researchers have only recently started to look at whether speakers also align in their co-speech gestures. So far, there is first empirical evidence for inter-speaker coordination of gestural behavior in both experimental settings (e.g. Holler & Wilkin 2011, Mol et al. 2012), as well as in natural dialogic interaction (Bergmann & Kopp 2012).

In the present study we aim to broaden the scope of this research by focusing on gestural alignment in human-computer interaction. Given that alignment in human-computer interaction has been shown to be even stronger than in human-human interaction (Branigan et al. 2010), we provide a first empirical test investigating in how far people converge gesturally with virtual humans. These embodied agents add human-like appearance and behavior to the interaction with computers to allow for natural and intuitive interaction with technical systems. At he same time, virtual humans provide a powerful tool to study inter-speaker coordination because the show flexible and controllable behavior, likewise.

The present study tests whether humans adapt their gestural behavior to an embodied, virtual agent. Participants were engaged in a communication game with a virtual agent in which tangram figures had to be explained to one another. In a between-subject design the virtual agent used different variations of gesturing behavior whereby handedness was systematically manipulated (two-handed vs. one-handed gestures vs. a control condition without any agent gestures). Each experimental condition consisted of 10 stimulus items provided by the agent and 10 response items to be described by the participant.

Our results from the analysis of 27 human-agent interactions demonstrate that participants indeed adapt their gesture use according to the agents' stimulus gestures in the experimental conditions. Remarkably, this alignment effect is particularly due to the fact that speakers' gesture adaptation does not occur immediately, i.e., not for the first items of the very experimental conditions, but rather for the recent items within the very conditions. Moreover, the alignment effect is stronger within participants than across agent and human. In other words, a speaker's consecutive gestures influence each other more than the gestures of the agent interlocutor (in terms of "self-alignment"), albeit the effectiveness of "other-alignment". And finally, there are obvious individual differences in participants' gestural adaptation. While some speakers show a strong tendency for coordinating their gestures' handedness, others did not deviate from their initial choice of handedness at all. In sum, we interpret these results such that there is an individual adaptation threshold which might depend on neural, cognitive, or social factors (cf. the "gesture threshold" in the GSA framework; Hostetter & Alibali 2008).

Improvised interactions dealing with machinery troubles

Masato Komuro, Mayumi Bono

The purpose of this research is to reveal a human actor's practices in dealing with troubles in the interaction between a robot and actors on the stage. Around 80 gigabytes of video data were recorded over six days of rehearsals (including a dress rehearsal) for a play called Night of the Galaxy Express, using high-definition cameras.

The Human-Robot Theater production Night of the Galaxy Express, directed by Oriza Hirata in collaboration with Osaka University, is unique in that it replaces a human actor with a robot. Night of the Galaxy Express is a very famous novel in Japan, written by Kenji Miyazawa. Oriza Hirata rewrote it as a stage play for human actors and directed it in France in 2011 and Japan in 2012. Then, in 2013, Night of the Galaxy Express was restaged as a "Robot Theater Project", in which a main character, "Campanella" was played by a robot called "Robovie" developed in the Ishiguro Laboratory at ATR Intelligent Robotics and Communication Laboratories, Osaka University.

In the role of Campanella, Robovie needed to interact naturally with human actors. However, various difficulties arose from the casting of the robot: for instance, unlike human actors, robots cannot improvise lines or adjust to time lags or mistakes by human actors, but instead merely utter preprogrammed lines at pre-determined times. Further, the robot

experienced many technical problems where it cached over and did not accept signals. As a result, it uttered lines at the wrong times or failed to utter them when required.

To address these troubles, the actors took various actions. In this research, we analyzed actorly behaviors to deal with troubles on the stage due to the robot's actions, focusing on multimodal features: utterances, gestures, gaze, nodding, and posture. We identified three types of actorly practices, implying different views of the problem. (1) Machinery Troubles: actors dealt with onstage issues simply as machinery troubles rather than interacting with the robot as with another person. (2) Interactional Troubles: actors dealt with machinery troubles as part of the onstage interaction, in front of the audience. (3) No Troubles: Troubles appeared neither Machinery nor Interactional.

Through these analyses, we revealed the actorly practices that maintained the coherence of the onstage interaction and progressed the action of the play. By improvising lines or gestures, the actors changed the addressee or sequence structures, bringing the script back on track in dynamic, multimodal ways. A theatrical performance constitutes a world of pre-established harmony: the actors know who the next speaker is and what he or she is going to talk about. Additionally, the performance of a play is obviously different from natural conversation in the real world (Sacks, Schegloff, & Jefferson, 1974). Nevertheless, the actors' improvisational ways of dealing with onstage troubles are "real" practices, in that they refer to the rules of natural conversation. Thus, focusing on these practices will help us reconsider the design of robots intended to interact with human beings.

Overall, the data show that gestures usually align with multiple lexical elements (agents and/or patients as NPs or pronouns in combination with the predicate, complements, conjunctions, etc.). As for the predictions, preliminary results indicate that gestures tend not to align with zero anaphora and pronouns, i.e. known/accessible referents, replicating previous findings. Observer-viewpoint gestures are more likely to align with lexical NPs (agents and patients alike) than character-viewpoint gestures. Character-viewpoint and observer-viewpoint gestures are equally likely to align with predicates plus agents or patients. However, observer-viewpoint gestures tend to include the agent, whereas character-viewpoint gestures tend to include the patient.

Deictic gesture, gaze and mobility. How a robotic museum guide initiates the transition to the next exhibit Karola Pitsch

Karola Pitsch

Considering "multimodality" as a basic condition of face-to-face communication points to its situated nature: Starting point for analysis are all those resources which are made relevant by the participants themselves during a conversation. In this vein, talk, gesture, gaze, body orientation and physical structures of the environment are closely interrelated and are best conceived of as complex "communicative shapes" or "contextual configurations" (Goodwin 2000). Investigating referential practices, studies have often focused on the interplay of talk and deictic gestures. This might be partly due to particular study set ups in the laboratory, in which participants are e.g. seated and/or already visually co-oriented to a particular location (e.g. de Ruiter et al. 2012). However, investigating participants interacting "in the wild" also the role of the speaker's gaze, bodily conduct and orientation in space becomes visible and how these are designed in concert with talk and gesture (e.g. Mondada 2009).

This situated nature becomes particularly evident when investigating guided museum tours. Guides do not only attempt to focus the visitors' attention on a particular object through pointing when standing in front of an exhibit (stationary phase), but they also attempt to re-orient them when attempting to make the group move from one exhibit to the next (mobile phase) (Stukenbrock 2012, Pitsch 2009). The complexity of the ways in which the different communicational resources are interrelated, becomes particularly visible once a technical system, e.g. a robot, attempts to assume the guide's role (e.g. Pitsch et al. 2013a, b).

In this talk, we investigate such moments of transition, in which a humonaid robot (Aldebaran NAO) assumes the role of a museum guide and attempts to move visitors in space from one exhibit to the next. The design of the robot's conduct when initiating these transitions involves shifting gaze, a deictic gesture, turning its body and walking to the next location. A particular difficulty arises from the circumstance that this robot, on the one hand, needs to shift its head to orient itself in the environment (and detect specific landmarks) as a prerequisite for walking in the right direction. On the other hand, the robot's head orientation in concert with the deictic gesture, shifting body orientation and walking movement serves as a display and communicational resource for the visitors.

Based on a corpus of video-recordings of human-robot-interaction with this robotic guide in a real-world museum, we address the following questions:

How do visitors interpret the robot's multimodal conduct? Which role does the deictic gesture play in concert with the other communicational resources' Analysis will show how uninformed visitors to the museum (mainly children between 4 and 10 years) react to the orienting function implied in the robot's multimodal practices. Implications will be drawn from the empirical analysis for further designing the robot's conduct and for understanding the multimodal interplay of communicational resources around deictic gestures. In this sense, human-robot-interaction becomes an empirical tool for investigating gesture and procedures of multimodal communication.

Alignment: Talks

Scales of gestural progression and information status in discourse

Sandra Debreslioska, Marianne Gullberg

Speech and gesture are known to work together to make discourse coherent and cohesive. For example, there is an increase of gesture use for new/less accessible discourse entities, but a decrease for given/more accessible ones (e.g. Gullberg, 2006; Levy & McNeill, 1992; McNeill & Levy, 1993). Information status and accessibility of referents also influences gestural means of representation. More accessible referents tend to be associated with character-viewpoint gestures whereas less accessible referents are associated with observer-viewpoint (Debreslioska, Özyórek, Gullberg & Perniss, 2013). Furthermore, McNeill (1992) has hypothesized that information status will influence 'gestural complexity' in a context where reference is maintained. This means that with increasing novelty there will be an increase in expressive complexity from zero anaphora and pronouns to noun phrases (NPs) in speech, and from no gesture and beats/pointing via observer- to character-viewpoint gestures. This study tested McNeill's (1992) hypothesis with a focus on viewpoint gestures deriving the following predictions:

1) Gestures will not align with zero anaphora and pronouns
2) Observer-viewpoint gestures will align with NPs
3) Character-viewpoint gestures will align with predicates

We elicited narratives from 20 native German speakers using a picture story (72 target pictures). Half of the target events were intransitive (i.e., involving only an agent) and half transitive (i.e., involving an agent and a patient). Agents were either animate or inanimate, whereas patients were always inanimate. The analysis focused on the exact temporal alignment between gestures and speech in clauses describing the target events. All gesture strokes were identified and coded for viewpoint (character- versus observer-viewpoint). NPs, pronouns and predicates were counted as aligned with a gesture if at least one syllable co-occurred with a gesture stroke.

The results broadly support McNeill's (1992) hypothesis for a link between speech and gesture at the discourse level with an increase in information novelty reflected in more linguistic material (e.g., NPs instead of pronouns), and a progression from no gesture to observer-viewpoint gestures. However, the progression from observer- to character-viewpoint gestures is less clear. Specifically, it remains unclear whether character-viewpoint gestures aligning with predicates+patients express more novel information than observer-viewpoint gestures aligning with agent+predicates. Therefore, we will argue for refinements of the gesture scale, and discuss implications for a view of the discursive speech-gesture relationship.

Hand-to-hand conflict & consensus: Gestural alignment in argumentative vs. affiliative conversations

Patricia Lichtenstein, Alexandra Paxton, Rick Dale

Gesture is an integral part of interpersonal communication, enabling speakers to complement, reinforce, and supplement speech in interesting and dynamic ways (e.g., Kendon, 1972, 2004; McNeill, 1992, 2005). In addition to increasing efficiency in language comprehension (Kelly et al., 2010) and fluency in speech production (Alibali et al., 2000), gesture helps to establish communicative common ground between interlocutors (Brennan, Galati, & Kuhlen, 2010). Common ground (i.e., the body of information shared between interlocutors) is considered essential to facilitating the process of interpersonal convergence, the tendency of interlocutors to become more similar in behavior and cognition as a result of their interaction (also known as alignment, coordination, or synchrony; Louwerse et al., 2012).

Gestural convergence has increasingly been a topic of interest among researchers (e.g., Bergmann & Kopp, 2012; Kimbara, 2008; Mol et al., 2012). However, in both convergence research generally and gestural convergence research specifically, the range of communicative contexts under consideration has been limited. Previous research tends to focus on convergence during positively valenced (Sebanz, Bekkering, & Knoblich, 2006) or affect'oneutral (e.g., task'obased interaction; Louwerse et al., 2012) communicative settings, but the space of human social experience is much broader. Emerging work within the convergence or synchrony literature is beginning to shed light on other communicative settings, including conflict (Paxton & Dale, in press).

Our study has three primary goals: (a) to characterize the differences in the use of gesture during argumentative versus affiliative conversations; (b) to determine whether and how gestural convergence (along temporal and categorical realms) appears during argumentative conversations, as compared with affiliative conversations; and (c) to investigate the relationship between the continuous dynamics of gestural movements and the categorical classifications of gesture, and examine how this relationship varies across discursive contexts. Toward these ends, we compare quantitative and qualitative features of gesture during argumentative and affiliative dyadic conversations. We adopt an annotation schemata reflective of our multilevel approach: we focus on functional dimension features such as iconicity and metaphoricity, formal dimension features such as trajectory and shape, and dynamic dimension features such as amplitude and velocity. Our discussion centers on situating gesture in the broader interpersonal convergence literature

Listener head gestures and the co-construction of narrative timing

Eric Pederson

This study examines on the use of head gestures for listener back-channeling during sustained narratives – focusing especially on the way interpersonal head gestures coordinate the temporal structure of a narrative.

Data is drawn from video materials of non-traditional story-telling episodes between largely mono-lingual Tamil speakers recorded in rural South India. These narratives were chosen because 1) many speech communities of India are known for their distinctive use of conventionalized head gestures (marking at least a three way emic contrast between nods, shakes, and wags) and 2) Tamil story-telling relies particularly heavily on listener participation. Also worth mentioning is the long-standing tradition of codified head gestures in classical dance forms (e.g., Bharatanatyam), etc. which suggests particular cultural emphasis on head gesturing.

The timing of the speakers' utterances (beginning of phrases, pauses, ends of phrases) was examined for correlations with listener behavior, most specifically listener back-channel vocalizations ("mm" "uh huh", etc.) and head gestures (type, duration, and amplitude).

There is remarkably close verbal synchronization between the two conversants: The beginning of a speaker's passage appears to best correlate with the end of back-channel vocalization, not with the end (nor amplitude) of listener head gesture.

Speaker head gestures are less common than speaker hand gestures and generally correlate with onset of listener head gesture. Unsurprisingly, speaker manual gestures best correlate with their apparent lexical affiliate in the speaker's narration. The timing of listener head gesture best correlates with either speaker gesture or gaze shift. It correlates only weakly/indirectly with the end of speaker vocalization. That is, interpersonal head gesture patterns appear to correlate highest with one another rather than with the other's vocalization. This demonstrates a relatively strong interdependence of head gesturing between speaker and hearer. Remarkably, however, there is relative independence between the head gesturing and speech across the speaker and listener.

This suggests a model of communication in which the interpersonal temporal control mechanisms of speech and head gestures across participants operate at least partially autonomously.

Experimental 3: Talks

The tone we dont hear - analysing the modifying functions of gestures

Farina Freigang, Stefan Kopp

Just as speech carries semantic propositions as well as modal and affective tones in prosody (e.g., Lu, Aubergöe, and Rilliard, 2012), there can be "tone" in gesture beyond propositional con- tent. Kendon (2004) differentiates between three pragmatic functions in gestures: The performative function carries the proposition of an utterance, the modal function implicates how a verbal utter- ance should be interpreted, and the parsing function contributes to the structuring of the utterance. Teöendorf (2005) investigated the Spanish "brushing aside" (annoying objects) gesture and found four pragmatic functions: turn taking organisation, structuring the verbal utterance, "metalinguis- tic" comment, and evoking/performing a speech act.

Here we are concerned with how gestures realise such modifying functions that operate on top of the propositional meaning. In particular, we are interested in how the evocation of these functions is restricted by the gesture itself while it also depends on the verbal and situational context. For example, a gesture holding up a fist in the air along with the words "she beat him badly" will be interpreted as depicting the action itself, while the same gesture along with an utterance like "I won the lottery" would be taken as expressing a feeling of joy. Yet, the very form of the gesture excludes other interpretations.

We hypothesize that certain sub-categories of modifying functions can be identified, which are often conveyed less explicitly and at lower degrees of intentionality (Allwood, 1976), i.e. they are mainly indicated (unconsciously sent) or displayed (consciously conveyed) but rarely signalled (intended to be recognised). We propose four distinct types of modifying functions: focusing, at- titudinal, epistemic, and emotional. (1) Focusing gestures signal relevance and importance of a proposition (e.g., "That was something!" accompanied by a pointing gesture into the air). (2) At- titudinal gestures reflect a stance towards the matter of a proposition, e.g., admiration, ignorance, doubt, disappointment, or contempt. Ignorance, e.g., can be displayed by a "throwing something over the shoulder" gesture on the words "they offer chips today". (3) Epistemic gestures display the degree of certainty of a proposition (e.g., the common wiggling with the hands to express un- certainty). (4) There may be emotional gestures displaying an affective state, e.g., anger, sadness, fear, boredom, or joy. Note that gestures can fall into more than one category at the same time or, depending on context, they may also shift in between these categories.

We have conducted an empirical pre-study on the occurrence of these gesture functions in the domain of describing "impossible objects" (optical illusions). Results show that modifying gestures (in this domain, predominantly focusing and epistemic) are an integral part of natural communication and occur frequently. A first systematic study is underway that investigates the interpretation of such gestures in isolation as compared to with different co-occurring verbal utterances. We will discuss methodological challenges and present first results.

Can you handle this' How object affordances determine representation technique in gesture

Ingrid Masson-Carro, Martijn Goudbeek, Emiel Krahmer

Previous research suggests that mental images are grounded in our perceptuo-motor experiences. One line of research within this view has focused on object affordances (i.e., action possibilities that objects allow for) as constituents of knowledge, revealing a tight link between the perception of affordances and language comprehension (Fischer & Zwaan, 2008). Affordance-related research has shown that merely viewing an object already triggers the activation of the motor processes associated with physically grasping that object (Ellis & Tucker, 2000), and of the particular handshape associated with executing its action (Bub & Masson, 2006). This has implications for gesture production, in that it suggests that perceiving objects with a manual affordance should evoke stronger motor simulations, which could in turn result in higher gesture rates (Hostetter & Alibali, 2008). Indeed, past studies have found more co-speech gestures when describing spatial or motoric tasks (Hostetter & Alibali, 2008), or have reported the production of more character viewpoint gestures in retelling events featuring an entity performing handling actions (Parrill, 2010). Beyond that, our knowledge of how object characteristics influence gestural patterns is scarce.

In this paper we look at action and gesture in a more subtle way, by looking at whether the passive observation of objects with an affordance activates simulations of proper object use, as reflected in speech-accompanying gestures. Crucial to our research is to elucidate what representation techniques are evoked by object affordances, that is, whether speakers use gestures to mime the object's conventional use ("handling gestures" -analogue to character viewpoint gestures), or produces gestures that depict static visual characteristics of the objects such as their shape, size or contour ("form gestures"). For this purpose, we collected a corpus of 1120 multimodal descriptions from 40 speakers about daily objects of two kinds, namely objects that have a concrete function to the hand (manual affordance group) such as a cheese slicer, and objects with no specific manual function (control group), such as a sink. The validity of the materials was assessed in a normative study evaluating object recognition, familiarity, usage frequency, functional and visual complexity.

We found that descriptions of objects with a manual affordance were more often accompanied by an iconic gesture than descriptions about the control objects. Preliminary analyses show that objects in the control group were represented mostly by gestures depicting their form (96%), with very few gestures exemplifying the object's purpose (4%). For objects with manual affordances, handling gestures (59%) dominated over form gestures (41%) ($p<.05$). Thus, our results suggest that the mere perception of manual affordances (without attending to action stimuli per se) is enough to influence whether speakers will produce gestures at all, and to determine the gesture representation technique, in line with previous behavioural studies, and as predicted by the GSA framework (Hostetter & Alibali, 2008). In addition, we aim to discuss data from more fine-grained analyses of form and handling presentations, uncovering general patterns and strategies in the communication about objects.

Iconicity in the generation of vocal conventions

Marcus Perlman, Rick Dale, Gary Lupyan

Conventional signed languages are generated from spontaneously created, motivated gestures such as pantomime and pointing (Armstrong & Wilcox, 2007). Yet, in theorizing about language evolution and glossogeny, it is often argued that the vocal modality does not afford the same opportunity for the creation of motivated gestures. Hence it is often reasoned that, in the vocal modality, "the issue of conventionalizing already meaningful acts never arises" (Tomasello, 2008: 228). We questioned this common assumption by performing a series of semiotics experiments examining the potential for speakers to create conventional signs in the vocal modality (cf. Galantucci & Garrod, 2010). The results show that, under some circumstances, nonverbal vocalizations can convey sufficiently precise information to ground the emergence of a spoken communication system even without prior conventionalization.

In an initial experiment, ten pairs of participants played an iterative "vocal charades" game. In the game, each player held a stack of twelve shuffled cards. Printed on each card was one of eighteen different words: attractive, ugly, bad, good, big, small, down, up, far, near, fast, slow, few, many, long, short, rough, and smooth. A given word was always mixed with its opposite in a player's stack; players held some words uniquely while others were held by both players. Over the course of ten rounds, players took turns vocalizing the words on their cards in an effort to get their partner to guess the word. Critically, they were not permitted to use words or gestures. Players had ten seconds to guess; correct guesses were immediately noted by the vocalizer, otherwise the correct word was noted at the end of the turn.

Praat was used to measure the acoustic properties of each sound, including its mean, minimum and maximum pitch, duration, intensity, and harmonics to noise ratio (HNR), as well as the number of sound repetitions. Partner's guesses were also analyzed. The results show that performance efficiency improved across the ten rounds. Both the number of sounds produced and the number of guesses per turn decreased over rounds, as the correct answer was guessed with fewer attempts based on fewer sounds. There is also evidence that the sounds became increasingly conventionalized over the course of the game, reflected in increasingly short duration and increased stability in form.

Additionally, iconicity appears to pervade the conventionalized forms that were developed for each of the eighteen words. Players reliably discovered sounds with similar acoustic properties for each particular word (e.g. rough had a lower higher harmonics to noise ratio and a higher intensity than smooth; fast had a higher intensity and more repetitions than slow). Logistic regression models conducted on all pairwise

word comparisons (i.e. each word compared to each other, not just opposites) showed that 147 of 154 comparisons had at least one of the five acoustic measures as a reliable predictor ($p < .001$ for each).

We conclude that the vocal modality affords greater potential for the creation of motivated, "already meaningful" gestures than commonly realized. The findings raise the possibility that, as with signed languages, conventionalized spoken language may be generated from an initially motivated system.

Parallel Session 9 – Friday, 11th 9:00 am - 10:40 am

Reference and multimodality in under-documented languages: Panel

Proximal pointing and person reference in Central Australian sand stories

Jennifer Green

Sand stories narrated by Indigenous women from Central Australia contain complex multimodal utterances that include speech, song, sign, gesture and semi-permanent inscriptions (Munn 1973; Wilkins 1997; Green 2009, 2014). These diverse semiotic resources form loosely co-ordinated dynamic partnerships or "ensemble" (Kendon 2004, 2008), exploiting the affordances of the various media employed. Actions with "graphic consequences" play an integral role (Goodwin 2003, 2007; Murphy 2005). Seated storytellers clear a drawing space on the soft ground in front of them, and the graphic schema that populate this space provide anchors or "targets" (Streeck 2011) for the unfolding narrative. In some stories leaves and other small objects are used to represent story characters, arranged in configurations somewhat like a miniature stage set. At the end of each scene the space is erased before the drawing begins again.

A sand story performance unfolds within the framework of several interlocking spatial fields, including the 3-dimensional space around the narrator and the ground space in front of them "where objects are visually available to both interlocutors" (Levinson & Wilkins 2006: 567). Deictic gestures are a key device used to meld speech and graphic imagery together. Pointing "laminates and transposes" (Haviland 2003: 158) these conceptual spaces and provides vectors of connection to the visible, the distant, and the imaginary. Green & Wilkins (in press) demonstrate how reference tracking in sand stories can be achieved, in the absence of speech, by complex utterances consisting of pointing, sand graphs, and conventionalised sign. Directional precision is achieved by pointing actions made with the hand, or augmented by the use of artefacts of the storytelling practice, such as sticks or wires.

Studies of pointing in Indigenous Australia (Haviland 1993, 2000; Levinson 1997; Wilkins 2003) show that pointing is socio-culturally complex. For example Wilkins (2003) examines cultural variation in Arrernte pointing, challenges the belief that particular pointing behaviours are universal, and places pointing within a structured system of semiotic signs. These studies typically draw attention to the ways that pointing is used within absolute geo-centred frames of reference. Although narrators of sand narratives certainly employ such large-scale gestures, in this paper I discuss the semiotic practice of pointing to proximal spaces and show how it plays a key role in maintaining person reference. This adds to previous analyses by focusing on spaces that are near-at-hand, and the ways that individuated visible targets are brought into communicative salience by close-range acts of pointing and touching (cf. Streeck 2009). I draw on an extensive corpus of videoed recordings of Arandic sand stories and of multimodal interactions in a Ngaatjatjarra children's kinship guessing game from the Western Desert of Australia. I suggest that the precision afforded by use of the hand as a close-up semiotic instrument allows more variation in pointing hand shapes than has been previously described for these languages. A better understanding of how reference is achieved and maintained requires attention to the interplay between both micro- and macro-scale deictic practices.

Unnecessary points: A study of ASL reference tracking

Anne Therese Frederiksen, Rachel I. Mayberry, Jennifer Green

In the human communicative domain, pointing is ubiquitous. For example, gesturers tend to establish and refer back to entities by indicating abstract locations in space. In signed languages, one can refer to persons or entities that are present physically or in the discourse with indexical or anaphoric points. Whether such points should be considered linguistic or gestural is under debate (Cormier et al, in press). Here we focus not on the linguistic nature of points in signed languages, but on their referential use in narratives.

In the human communicative domain, pointing is ubiquitous. For example, gesturers tend to establish and refer back to entities by indicating abstract locations in space. In signed languages, one can refer to persons or entities that are present physically or in the discourse with indexical or anaphoric points. Whether such points should be considered linguistic or gestural is under debate (Cormier et al, in press). Here we focus not on the linguistic nature of points in signed languages, but on their referential use in narratives.

In a study of reference in narratives, Swabey (2002) found a low number of pronominal points in native ASL (0.6%) compared to pronouns in English (26%). This is somewhat surprising. Because ASL has pronouns, we would expect them to pattern like many spoken languages where pronouns are the primary means for referring to previously introduced discourse entities.

We found that native signers used anaphoric points only 1% of the time. Agreement verbs constituted 15% of the maintained references, nouns 7%, constructed speech/action 16%, and zero anaphor 39%. These results suggest that native ASL signers substantially disprefer anaphoric points for

maintaining referents. Although agreement verbs occurred more frequently than did anaphoric pronouns, we found only 1 instance of a signer using two locations in space for verb agreement. Instead signers preferred a strategy involving role shifting for signaling verb agreement. An agreement verb can begin or end at the signer's locus, but by role-shifting (turning head/torso towards a locus) the signer signals to the addressee that the first person reference has shifted and now indexes a referent from the narration, rather than the signer herself. It is arguable whether verb agreement of this kind should be considered pronominal.

Our findings suggest that anaphoric points (of any type) do not play the same prominent role in reference tracking in ASL as in spoken languages. Instead, referent maintenance is achieved primarily by classifiers and zero anaphora. Whether ASL signers' avoidance of pronominal points is limited to certain types of discourse remains an empirical question, as does the reason for the avoidance.

Gesture-speech synchronization and the demonstrative-definite referential distinction in arapaho

Rich Sandoval

Gesture can play a role in distinguishing different types of linguistic reference. Research in this area focuses on the unique referential features that gesture brings to otherwise spoken utterances (e.g. Enfield, Kita, and De Ruiter 2007; Wilkins 2003). Less understood are referential features that are dependent on the entire gesture-speech ensemble, which I address through research on Arapaho (a Native American language).

A key issue for any linguistic description is if/how noun phrases [NPs] signal the distinction between demonstrative and definite reference [DEM-DEF distinction]. For human interaction, the two types of reference are fundamental: In demonstrative reference, a referent is clarified for other interactants, as a means to develop common ground; in definite reference, a referent that is already familiar amongst interactants is identified as such, as a display of common ground (cf. Hanks 1990; Clark 1996). Thus, a marked DEM-DEF distinction enables interactants to socioculturally fine-tune basic acts of information sharing (cf. Enfield 2006). Cross-linguistically, English is typical, marking the DEM-DEF distinction with determiners (i.e. definite, "the", and demonstratives, such as "this" in "look at this spider"). From a speech-only perspective, Arapaho is notable for not marking the DEM-DEF distinction in NPs: A set of determiners (e.g. nehe") each ambiguously signals both types (e.g. nehe" wox "the/this bear") (Cowell and Moss Sr. 2008). There is strong analytical motivation, however, to go beyond speech-only: Arapaho gesture is highly conventional, including absolute pointing and features from an alternate sign language. Using the Arapaho Conversational Database (2011), a video-based interactional corpus, I observe that NPs produced through speech are regularly indexed through gesture. For NPs with determiners, I argue that gestural pointing and speech are synchronized together in different ways as a means to mark the DEM-DEF distinction.

A key issue for any linguistic description is if/how noun phrases [NPs] signal the distinction between demonstrative and definite reference [DEM-DEF distinction]. For human interaction, the two types of reference are fundamental: In demonstrative reference, a referent is clarified for other interactants, as a means to develop common ground; in definite reference, a referent that is already familiar amongst interactants is identified as such, as a display of common ground (cf. Hanks 1990; Clark 1996). Thus, a marked DEM-DEF distinction enables interactants to socioculturally fine-tune basic acts of information sharing (cf. Enfield 2006). Cross-linguistically, English is typical, marking the DEM-DEF distinction with determiners (i.e. definite, "the", and demonstratives, such as "this" in "look at this spider"). From a speech-only perspective, Arapaho is notable for not marking the DEM-DEF distinction in NPs: A set of determiners (e.g. nehe") each ambiguously signals both types (e.g. nehe" wox "the/this bear") (Cowell and Moss Sr. 2008). There is strong analytical motivation, however, to go beyond speech-only: Arapaho gesture is highly conventional, including absolute pointing and features from an alternate sign language. Using the Arapaho Conversational Database (2011), a video-based interactional corpus, I observe that NPs produced through speech are regularly indexed through gesture. For NPs with determiners, I argue that gestural pointing and speech are synchronized together in different ways as a means to mark the DEM-DEF distinction.

Manual and non-manual pointing as alternative formulations of place reference in Kula

Nicholas Williams

This paper explores the use of pointing practices in the context of place reference among speakers of Kula (Timor-Alor-Pantar, Indonesia). The analysis focuses on the use of pointing gestures as alternative formulations of place reference, occurring either independently or along with other verbal formulations. This work builds on previous conversation analysis (CA) work on person reference. The domain of place reference has just begun to be explored from in CA (Enfield 2013).

The topic of reference is one of foundational importance in the conversation analytic literature. However, despite early work by Schegloff (1972) on place reference, the focus has been primarily on reference to persons. Work on person reference established several basic principles of conversational organization, including the notion of "preference". Preferences for "recipient design" and "minimization", which are "found widely operative in conversation", were first shown to operate in the domain of person reference (Sacks & Schegloff 1979). Work on person reference has recently been extended to lan-

guages other than American English (e.g. Stivers & Enfield 2007) but remains a "sketch of a sketch" of the full "systematic organization of person reference resources" (Schegloff 1996).

The topic of reference is one of foundational importance in the conversation analytic literature. However, despite early work by Schegloff (1972) on place reference, the focus has been primarily on reference to persons. Work on person reference established several basic principles of conversational organization, including the notion of "preference". Preferences for "recipient design" and "minimization", which are "found widely operative in conversation", were first shown to operate in the domain of person reference (Sacks & Schegloff 1979). Work on person reference has recently been extended to languages other than American English (e.g. Stivers & Enfield 2007) but remains a "sketch of a sketch" of the full "systematic organization of person reference resources" (Schegloff 1996).

How these findings relate to the organization of place reference remains unknown. The bulk of work on place reference has been from a descriptive/typological, and/or cognitive/psycholinguistic nature (e.g. Senft 1997, Anderson & Keenan 1985, Levinson & Wilkins 2006, among many others). Recent calls have been made for extending the coverage of research on reference to include a typologically diverse set of languages and a broader set of reference domains, including place, time, object (Enfield 2013).

This paper presents preliminary findings of work documenting practices of place reference among speakers of Kula, a non-Austronesian language of Indonesia. Data is drawn from recordings of natural interaction in everyday settings. Frequently used options for formulation of place reference include place names, descriptions associated with particular persons, demonstratives, elevationals ("down there", "up here", etc.), and pointing gestures. Pointing gestures are unique in that they can be deployed simultaneously with verbal formulations of place reference.

This paper analyses two types of points, 1) arm/index finger points and 2) non-manual, head/eye/lip points (cf. Enfield 2001). Kula speakers frequently employ manual and non-manual pointing in their references to places in the immediate surrounding. Points are absolute, directed to the existing place. Non-manual points ("S-points," Enfield 2009) are hypothesized to function as a minimal way to increment place references in two contexts: 1) when a problem of recognition arises with the first formulation and 2) to make an initial formulation more recognizable in a minimal way. I draw heavily on CA studies of person reference showing that 1) person references are preferably as minimal as possible (e.g. a pronoun is more minimal than a name) while maintaining recognition, (e.g. a name as more recognitional than a pronoun in initial position) and 2) the form of reference is sensitive to its sequential position (i.e. initial or subsequent). Other relevant aspects of pointing in place reference formulation in Kula are described as well.

Repair: Talks

Repair sequences in cross-signing

KangSuk Byun, Anastasia Bradford, Connie De Vos, Stephen C. Levinson, Ulrike Zeshan

The study reported here involves communication between deaf sign language users with highly divergent linguistic backgrounds who have no signed or written language in common. Unlike the semi-conventionalised contact language International Sign (e.g. Supalla & Webb 1995), we look at the earliest, least conventionalised stages of improvised communication, called "cross-signing". Our interest lies in the shared conversational infrastructure, as well as metalinguistic abilities, that allow signers to co-construct meaning across linguistic and cultural boundaries in this type of ad hoc communication.

Our data set consists of the first encounters between three dyads of signers of Korean Sign Language, Sign Language of the Netherlands, and Russian Sign Language, Hong Kong Sign Language, (totalling 60 minutes of signed video data). We here focus on Other-Initiated Repair (OIR) sequences that target the use of novel signs, a three-turn structure including the problem source turn (T-1), the initiation of repair (T0) and the turn offering a problem solution (T+1) (Dingemanse et al. 2013). Ongoing analyses have identified 50 OIR sequences in our data set.

We find that in most cases of T-1, signers use repetition, gestural holds, prosodic lenghtening and eye gaze at the addressee as try markers (cf. Moerman 1988). These try markers make relevant a contingent response from the interlocutor such as a nod indicating recognition. In some cases, the absence of backchanneling, also resulted in a problem solution being offered. Overall, OIRs were twice as likely to be preceded by a T-1 with try marking, than one without, suggesting that sign-producers may frequently anticipate trouble. This shows that, via try-marking, sign-producers might actively mobilize an OIR which is an otherwise marked turn.

Sign language users may face communicative problems that arise from the absence of a conventional language and are thus specifically associated with cross-signing. To resolve this communicative problem, signers capitalise on repair: a sequential infrastructure that is accessible to all, partially independent of language (Levinson 2013). Repair sequences are central to understanding the cooperative process of language creation in cross-signing settings. At T0, addressees frequently responded by repeating the sign that is the problem source, thus initiating restrictive repair. In the absence of linguistic convention, signers then use a wide range of semiotic resources to resolve reference at T+1: including logical inference, iconic depiction, and paraphrase.

One general consideration arising from these data sets involves the role of meta-linguistic skills. Preliminary findings

show variation in both the success rate in resolving reference and the diversity of metalinguistic structures that are used. It remains to be investigated how this may correlate with individual backgrounds such as age of sign acquisition, being fluent in multiple sign languages, and having international deaf social networks.

Assessing group synchrony during a rhythmic social activity: A systemic approach

Tariq Iqbal, Laurel Riek

Group interaction is an important part of human social behavior and an active area of research. During some group events, the actions performed by each member continually influence the activities of other group members]. This process of influence, or joint action, can create a state of interdependence, where each member's actions are determined by the actions of others, which leads to synchronized group activity. Group level synchrony may be an important behavioral indicator of group level cohesiveness, and also an important aspect of accurately understanding the affective behavior of a group.

We present a systemic method to automatically detect psychomotor group synchrony which incorporates multiple types of discrete, task-level events of individual group members. We extend the event synchronization method proposed by Quian Quiroga et al., as well as the follow-on work by Varni et al.

We employ the following steps to measure overall group synchrony:

1. Express all the events associated with each member over time as a time series.

2. Measure the pair-wise synchronization index (PSync) for each pair of members, while taking all types of events into account together.

3. Build a connectivity graph (CG) of the group based on the PSync values.

4. Calculate the individual synchronization index (ISync) of each member to the group from PSync and CG.

5. Measure the overall group synchrony, presented with the group synchronization index (GSync), from ISync and CG.

To validate our method, we designed an experiment where a group of four people played the Cup Game, a popular rhythmic game. In the game, each player stands at a table in a circle and taps, claps, hits, and passes their cup to others. The goal is for the group to be synchronous. During the game, we recorded video and depth data (i.e., skeletal movements) using two time-synchronized Microsoft Kinect sensors.

Different steps (i.e., clapping, moving or passing the cups) of this game were considered as different types of high level events, which were detected by combining the hand gestures of each player and their cup movements. We detected hand gestures using the skeletal data, and tracked the cup movements from the RGB video using standard computer vision techniques (i.e., blob tracking). We measured the GSync using these events.

A total of 22 people (50% female) participated in six experimental sessions, consisting of a group of four players in each session. Each group played two games. After each session, each player completed a questionnaire which asked them to rate which one of the two games they felt was more synchronous. Results suggest that the GSync values produced by our method agreed with the perception of the majority of participants across all six sessions.

Our method is useful to other researchers interested in automatically and efficiently exploring the plethora of information that a social group activity contains. In addition, our method is extensible to apply to any group activity where events are detectable as well as usable within multimodal systems.

Gesture as a resource for initiating repair

Kristin Mortensen

In recent years, the inclusion of bodily and material resources has increased dramatically in empirical micro-studies of social interaction. Most of them draw on ethnomethodology/ conversation analysis methodology in combination with findings from neighboring fields most noticeably context analysis and gesture studies (e.g., Streeck, Goodwin & LeBaron, 2011). Together they argue that participants draw on a range of resources from various "semiotic fields" (Goodwin, 2000) in the construction of social action. As a result, these studies argue that in adopting an emic perspective, so too must the analyst consider the range of possibly relevant resources in the participants' perceptual fields rather than a priori foreground one set of interactional components - such as gesture or speech (e.g., Hazel, Mortensen & Rasmussen, forthc.). Critics of such an approach to social interaction have argued that the empirical description of bodily conduct has -to date- not been described as systematic and recognizable aspects of interaction. For instance, Schegloff (2007: 11) argues that "there is no reliable empirical basis for treating physically realized actions as being in principle organized in adjacency pair terms". More recently, this view has been challenged by analyzing how bodily conduct, in the absence of speech, is organized in much the same way as has been described as first and second pair parts. For instance, a few studies have described how a bodily action is oriented to as an accountable first (e.g. Mortensen, 2012; Seo & Koshik, 2010) or second pair part (e.g. de Stefani & Gazin, in press; Arminen et al., forthc.), respectively.

The present paper adds to the description of how bodily conduct can perform an initiating action. The paper analyses how the body can serve as a resource for initiating repair. It describes how a hand gesture, "cupping the hand behind the ear", in the absence of vocal and verbal conduct is oriented to as a repair initiation in a foreign language classroom. The gesture, although it verbally does not indicate the type of trouble-hearing or understanding- the participant is having is treated as a hearing problem, and is generally followed by

a repeat of what is treated as the trouble source - the prior turn. However, the gesture is generally found in a sequential environment, in which the trouble source turn is somehow sequentially or "topically" problematic. In this way, the gesture functions similarly to "open class repair initiations" as described by Drew (1997) through verbal conduct.

Co-Speech: Talks

Co-thought and co-speech gestures are generated by the same action generation process

Mingyuan Chu, Sotaro Kita

People often spontaneously gesture when they speak (co-speech gesture). Some hypotheses claim that co-speech gesture and speech systems are one unity and inseparable (McNeill & Ducan, 2000; McNeill, 2005, 2012). Other hypotheses proposed that co-speech gestures are generated from the representational use of the action generation process (LeBaron & Streek, 2000; Kita, 2000; Kita & 'ozy'orek, 2003). In addition to co-speech gestures, people also spontaneously produce hand movements when they solve problems without speaking (co-thought gesture; Chu & Kita, 2008, 2011; Hegarty et al., 2005; Schwartz & Black, 1996), which is presumably generated from the representational use of the action generation process.

The current study aimed to contrast the above two hypotheses by examining the relationship between the frequency of co-speech and co-thought gestures (Experiment 1) and by examining the effect of object affordance (i.e., the possibility of action upon an object) on the frequency of two types of gestures (Experiment 2 and 3). If both types of gestures are generated from the representational use of the action generation process, the frequency of these two types of gestures should be positively related and should decrease when the stimulus object was less likely to be act upon.

The current study aimed to contrast the above two hypotheses by examining the relationship between the frequency of co-speech and co-thought gestures (Experiment 1) and by examining the effect of object affordance (i.e., the possibility of action upon an object) on the frequency of two types of gestures (Experiment 2 and 3). If both types of gestures are generated from the representational use of the action generation process, the frequency of these two types of gestures should be positively related and should decrease when the stimulus object was less likely to be act upon.

Habituating cognition: On the role of practice and the built environment in shaping spatial cognitive styles

Melanie McComsey

The spatial co-speech gestural practices of bilinguals who speak languages that differ in their spatial semantic typology can offer special insight into the classic problem of the relationship between linguistic code and cognitive style. While some previous studies have found evidence for "hybrid" gestural systems in bilinguals or for cross-linguistic transfer (e.g. Brown & Gullberg 2008, Cadierno & Ruiz 2006, Choi & Lantolf 2008, Nicoladis et al. 2007), others have found evidence for highly language-specific systems (e.g. Navarro & Nicoladis 2005, Neguerela et al. 2004, 'oz'oal'o'okan & Slobin 2000), even in very young bilingual children (Nicoladis et al. 2010), signifying separate conceptual systems for each language. What has been lacking to date, however, in studies of spatial language and cognition in bilinguals, is sufficient attention to the contexts in which bilingual speakers use their two languages, and how these contextual features might affect spatial cognition. The present study addresses this need by combining semi-experimental tasks to elicit spatial frame of reference speech and gesture with analysis of naturalistic spatial speech and gesture collected during long-term ethnographic fieldwork. Specifically, it investigates the spatial speaking and thinking patterns of four bilingual (Spanish/Zapotec) children who live in Juchit'on, Oaxaca, Mexico. In Juchit'on, rapid modernization has led to changes in the built environment and in related spatial practices, as well as to a growing prevalence of Spanish in what was historically a primarily Zapotec-speaking city. In attending to children who differ in their bilingual proficiency and in their typical spatial practices, this study suggests that ways of thinking are less coupled to specific language codes in Juchit'on than integrated into a complex system of communication that is locally specific. Typically co-speech gestures related to spatial frames of reference have been considered in situations where there is little within-population variation in cognitive styles and corresponding gestural practices. Juchit'on offers a unique case in which language contact as well as rapid change in the built environment and cultures of spatial practice seem to be related to variation in styles of conceptualizing space as evidenced through gesture.

Gesture serves the speaker more than the listener in descriptions of ineffable shapes

Zed Sevcikova Sehyr, Karen Emmorey

We investigated the role of co-speech gesture when speakers describe ineffable shapes that cannot be easily labeled (Atteneave and Arnoult 1956). Gesture is hypothesized to help express analog information not easily encoded in speech (McNeill 2000), to aid in lexical retrieval- particularly for spatial language (Krauss, Chen et al. 2000; Hostetter, Alibali et al. 2007), and to organize complex information for the speaker (Kita 2000). These functions point to an important role for gesture when speakers must uniquely identify non-nameable shapes. To investigate this role, 10 pairs of English speakers (N=20, mean age 24.9, SD = 5.3; 11F) completed

a referential communication task based on Clark and Wilkes-Gibbs (1986). Participants sat beside each other at a table, with identical twelve cards laid out in front of them. Each card contained one Atteneave shape. A low divider allowed the participants to view each other, but not the other persons' cards. For one participant, 'Director', the cards were pre-arranged in two rows of six shapes. For the other participant, 'Matcher', the same cards were randomly arranged. The task required the Matcher to re-arrange their cards to match the Director's layout based on the Director's spoken descriptions of the shapes. Participants performed the task six times, with the same cards re-ordered for each trial. Referring expressions were coded as shape-based (e.g. a rectangle with a triangle cut-out) or as lexical (e.g. looks like a mountain). Analyses examined representational gestures (iconic gestures bearing some resemblance to the shape being described). Preliminary results show a linear decline in iconic gestures across the six trials, $F(1, 7) = 6.23$, $p < 0.05$. 40% of all gestures were produced in the first trial compared with only 10% in the final trial. Gesture strings were more prevalent in the first than the last trial, and the duration of gestures decreased across trials. The length of spoken descriptions also decreased dramatically across trials: 281 sec (SD = 176) for trial one vs. 46 sec (SD = 25) for trial six, $F(1, 8) = 30.17$, $p < .05$. By trial three, speakers had switched from lengthy shape-based descriptions to lexical labels. There was a correlation between gesture use and spoken shape-based descriptions ($r2 = .35$, $p < .001$) and a negative correlation between gesture and lexical labels ($r2 = .23$, $p < .05$). These results suggest that shape-based descriptions favor iconic gestures and are consistent with previous studies on spatial language and gesture (Rauscher, Krauss et al. 1996). Further, the majority of gestures referred to parts of the shape rather than to the shape as a whole and closely followed the spoken part descriptions gestures dropped off as soon as speakers landed on a lexical label. Surprisingly, Matchers looked primarily at their cards and rarely at the Directors' gestures: Matchers directly observed only 5% of Directors' gestures, indicating that gestures had little or no communicative function. This striking result suggests that co-speech gesture functioned primarily to benefit speakers, aiding retrieval and packaging of spatial information for these shapes.

When anyone's listening: Automatization and gesture reduction

Prakaiwan Vajrabhaya, Eric Pederson

Studies have shown that when words are repeated in a conversation, they reduce phonologically; that is, subsequent mentions of repeated referents are shorter than the first mention (Lam & Watson, 2010; Baker & Bradlow, 2009; Aylett & Turk, 2004; Anderson & Howarth, 2002; Bard et al., 2000). Two models have been proposed to account for the reduction phenomena. The speaker-based model suggests that speakers reduce repeated articulatory sequences due to automatization (Bybee, 2002). On the other hand, the listener-based model suggests that speakers reduce only units referring to referents the listener already knows (Fowler, 1988; Fowler & Housum, 1987). More importantly, reduction of repeated referents is not only restricted to the phonological domain; there is evidence that it also occurs in co-speech gesture. For instance, Gerwing and Bavelas (2004) showed that gestures are smaller in size and precision is decreased when referring to old information; furthermore, Hoetjes et al. (2011) reported a qualitative and quantitative decrease in gesture when information is old. Thus, reduction of expressions referring to previously mentioned referents appears to be a systemic cross modal tendency.

In this study, participants watched a short demonstration of pizza making. They, then, described the procedure to confederates representing both a new and a repeated listener in a condition sequence of: (1) Listener A; (2) Listener B; (3) Listener A again (Galati & Brennan, 2009). This experimental design creates a paradigm where information is old to the speaker but new to the listener (condition 2) and where information is old to both the speaker and the listener (conditon3). This design allows us to determine whether any reduction is due to speaker-automatization or listener-sensitivity.

Our results show a reduction in the scale of co-speech gesture as well as a shift in type across both retellings regardless of whether the information is new or old to the listener (condition 2 and 3). Gesture was coded when it co-occurred with descriptions of the same event across the three conditions. Scale, measured by number of body parts in motion, statistically proved to reduce evenly across the three conditions. We also find a type shift in which gestures typically reduce from having available semantic content to beats or no gesture. In sum, when the same story is repeated, regardless of who is listening, gesture reduces quantitatively as well as qualitatively.

Our results show a reduction in the scale of co-speech gesture as well as a shift in type across both retellings regardless of whether the information is new or old to the listener (condition 2 and 3). Gesture was coded when it co-occurred with descriptions of the same event across the three conditions. Scale, measured by number of body parts in motion, statistically proved to reduce evenly across the three conditions. We also find a type shift in which gestures typically reduce from having available semantic content to beats or no gesture. In sum, when the same story is repeated, regardless of who is listening, gesture reduces quantitatively as well as qualitatively.

Space and viewpoint 3 : Talks

The emergence of linguistic use of space in an experimental gesture communication study

Emily Carrigan, Marie Coppola, Whitney Tabor

”This work employs experimental semiotics methodology to directly test hypotheses about the mechanisms underlying language emergence. In particular, we explore how space comes to be used to express argument structure, as in many established sign languages1. In spatial verb agreement systems, arguments are localized in space or on the body, and the verb's movement between those locations indicates arguments' thematic or grammatical roles. The consistent use of this device has been demonstrated in second- but not first-cohort users of Nicaraguan Sign Language (NSL2); however, no definitive account of its emergence has been provided.

We propose that the abstract use of space for argument structure develops from a “distalization” of embodied representations (cf. Frishberg3). The expression of certain types of events (e.g. semantically reversible events) begins with a fully embodied, mimetic representation, and both lexical and grammatical content are increasingly condensed in size (referential/role shift) and transferred to the hands (abstract use of space).

We propose that the abstract use of space for argument structure develops from a “distalization” of embodied representations (cf. Frishberg3). The expression of certain types of events (e.g. semantically reversible events) begins with a fully embodied, mimetic representation, and both lexical and grammatical content are increasingly condensed in size (referential/role shift) and transferred to the hands (abstract use of space).

In this pilot study, three pairs of hearing non-signers viewed videos of simple two-argument, semantically reversible events (like “a man taps a woman”), then “acted out” these events to a na'ove partner, while standing, using only their hands and body. Receivers demonstrated comprehension by selecting a matching picture. After incorrect choices, the producer re-described the event. Interactions were video-recorded, and coded offline for how producers designated characters' roles (e.g. gesture-order, embodied action, spatial modulation/movement of manual gestures).

In this pilot study, three pairs of hearing non-signers viewed videos of simple two-argument, semantically reversible events (like “a man taps a woman”), then “acted out” these events to a na'ove partner, while standing, using only their hands and body. Receivers demonstrated comprehension by selecting a matching picture. After incorrect choices, the producer re-described the event. Interactions were video-recorded, and coded offline for how producers designated characters' roles (e.g. gesture-order, embodied action, spatial modulation/movement of manual gestures).

Results suggest that space is an option for representing argument structure, but individuals vary in their use of this device. For example, in one strategy, producers embodied (acted out) the roles of both the agent and patient. In another male-female pair, each member of the pair embodied the action of the character matching their own gender, and manipulated the body of their interlocutor to embody the role of the other character (videos always featured one male and one female). In both strategies, agents and patients were associated with distinct non-neutral spatial locations, and gestures for verbs moved between these locations.

Results suggest that space is an option for representing argument structure, but individuals vary in their use of this device. For example, in one strategy, producers embodied (acted out) the roles of both the agent and patient. In another male-female pair, each member of the pair embodied the action of the character matching their own gender, and manipulated the body of their interlocutor to embody the role of the other character (videos always featured one male and one female). In both strategies, agents and patients were associated with distinct non-neutral spatial locations, and gestures for verbs moved between these locations.

Linguistic, gestural, and cinematographic viewpoint: An analysis of ASL and English narrative

David Quinto-Pozos, Fey Parrill, Stec Kashmiri, Sebastian Rimehaug

Studies of multimodal language make frequent use of video stimuli, but how does the cinematography used to present a sequence of events affect gestural or signed communication'o Do speakers or signers replicate the visual perspective of video stimuli in language or gesture because those features are part of mental representations formed during encoding'o Alternately, do properties of the language used (spoken or signed) encourage particular perspectives in the narrators' accounts of the events'o For example, some researchers have suggested that the use of the passive in ASL is correlated with differences in perspective (e.g., see Janzen et al., 2001), which may constrain signers.

In a recent analysis of a Brazilian Sign Language narrative, McCleary and Viotti (2010) note that a signer used linguistic and gestural resources to replicate a cinematographic shift of perspective. In our study, we ask whether this pattern can be reliably elicited in multiple English-speaking gesturers and ASL signers.

Eighteen English speakers (hereafter, speaker-gesturers) and ten ASL signers watched a video stimulus in which the perspective shifts between two characters across multiple scenes; one character is inside a room and the other outside and the video depicts camera shots from both locations. Participants were asked to describe the events depicted in the video (either in English or in ASL). Descriptions of three scenes (representing different cinematographic shots) were coded for linguistic properties and non-linguistic gesture.

Eighteen English speakers (hereafter, speaker-gesturers) and ten ASL signers watched a video stimulus in which the perspective shifts between two characters across multiple scenes; one character is inside a room and the other outside and the video depicts camera shots from both locations. Participants were asked to describe the events depicted in the video (either in English or in ASL). Descriptions of three scenes (representing different cinematographic shots) were

coded for linguistic properties and non-linguistic gesture.

We found that both groups of language users tended to show the events from a single perspective, that of the character: 67% of iconic co-speech gestures and 60% of non-lexical signs (i.e., use of constructed action by signers) showed action from the character's point of view. We suggest this is likely due to the nature of the events in the scenes. Neither group replicated the perspective shift of the stimulus. Instances of multiple simultaneous points of view were rare for both groups. Overall, we find that both groups used space in ways that were consistent with the character's point of view, rather than the viewer's.

We suggest that speaker-gestures and signers choose strategies that minimize cognitive load while maximizing visual detail of construal. It is more efficient to explain a scene from a character's perspective rather than adopting the perspective that the narrator experiences when viewing a scene.

Ongoing analyses are exploring the linguistic constructions that co-occurred (either simultaneously or sequentially) with the subjects' productions in order to determine if active or passive syntactic constructions can provide justification for the subjects' communicative choices.

Development of perspective choice and change in Japanese Sign Language

Yufuko Takashima

This paper examines development of perspective choice and change in deaf children‘s predicates in Japanese Sign Language (JSL). We analyze spatial expressions of deaf children in a private deaf school, educated in a bilingual program with JSL as a natural language and written Japanese. Using McNeil (1992)‘s animation story retelling task, we collected motion event predicates from 10 younger children (age 5-7) and 7 older children (age 10-12) learning in the school compared with 24 adults signers‘ utterance. These adults are all deaf from the birth and use Japanese Sign Language as their first language including native signers, early learners, and some late learners.

It has been investigated that the viewpoint indicated in child‘s co-speech gesture is different from its adult‘s in gesture studies of spoken language, focusing on the contrast between character and observer viewpoint. Adults tend to take the observer viewpoint while children are more likely to take the character‘s viewpoint. This shows children are not good at taking the observer‘s objective viewpoint. In signed language studies, on the other hand, some researchers have tackled with the use of depicting constructions, also called classifier constructions, related to the viewpoint matter. Cormier et al. (2013) compares deaf adults and children‘s utterances in British Sign Language and suggests adults‘ predicates containing depictive features are more complex than children‘s. This study shows it is difficult to learn how to use depictive construction because they should choose from several choices, such as character/observer viewpoints, size, and path or manner of motion.

Now we focus on change of perspective, called reference shift, and examine when signers change the viewpoint in an utterance. Although it seems easy to choose a fixed viewpoint, both adult and child signers in JSL show the change of the perspective during an event sequence.

Adult signers show strong tendency to take the character viewpoint and they frequently change to the other character's viewpoint. One of the reason is many lexicons of JSL need to be described on and from the first person's body. They seldom take an observer‘s objective viewpoint. They sometimes predicate the same scene twice sequentially one to the other characters‘ viewpoints to clarify whose viewpoint they are expressing from.

In contrast to the adults, younger signers show depictive predicates from both the character and the observer viewpoint. Unlike adults, they expressed an event from the camera angle. This indicates that they are not skilled on change of perspective and they express what they see with no change. In other cases, many children successfully take the character viewpoint as adults take because the lexicon need to take the perspective mandatory or they cannot describe the event.

Unlike hearing children, we found the signing children tend to take the observer perspective even adults do not use. This, however, does not necessarily means signing children are good at taking observer‘s objective perspective. They should learn perspective change under the limitation of the linguistic structures in the signed language unlike spoken language and we suppose that their immature ability to take the others viewpoint shows the difference from adults‘ expressions.

Parallel Session 10 – Friday, 11th 11:00 am - 12:40 pm

From gesture to word: Emerging sign: Panel

A scherzetto for four hands: assisted acquisition of Z

John Haviland

In a single extended family in a Tzotzil (Mayan) speaking community in the highlands of Chiapas, Mexico, a first generation sign language isolate is emerging with no input from other deaf people and no contact with any other sign language. Dubbed Z (Zinacantec Family Homesign, Chiapas, Mexico), the sign system is the principal means of communication between three deaf siblings and their hearing sister and a coeval hearing niece, along with a hearing nephew who represents a possible second generation of signers. Z shows clear links with the gestures of Tzotzil speakers, and it exhibits various kinds of nascent linguistic structure, from phonological and phonotactic systematization and morphological productivity, to iconic strategies in lexicalization, emerging grammatical categories, constituent structure, and even an insistent set of language ideologies.

Research has documented the acquisition of both spoken Tzotzil and signed Z by the youngest member of the signing community, now six years old, from his earliest linguistic productions starting when he was 11 months of age. The progression has been discontinuous, as a result of family circumstances that periodically separated the child from his deaf mother because of fears that he might not "learn to speak."Nonetheless, the miniature speech community of the deaf signers also took explicit and insistent steps to insure that he would also learn to sign.

This study concentrates on the evidence for both emerging linguistic structure and conscious metalinguistic ideology in several videorecorded interactions in which the young child is explicitly trained in signing by his deaf uncle.

Repair initiation in gesture and sign: A study of Chatino Sign Language

Kate Mesh, Lynn Y. S. Hou

This paper considers interaction in Chatino Sign Language (hereafter CSL), a developing language used by 10 deaf individuals and their hearing family members in Oaxaca, Mexico ([8]). Video recordings of interaction in the language reveal an unusual communicative circumstance: extended family members of deaf individuals may know their relatives very well, sharing with them extensive knowledge of the family and larger community, and nevertheless control little of the signed language that is the sole medium of communication for the deaf. Despite great potential for communicative difficulty, deaf signers and their hearing relatives with varying signing skills successfully exchange complex messages using manual and non-manual signals, and effectively resolve communicative troubles during unfolding interaction.

We investigate the manual and non-manual gestures/signs used by the signers for other-initiated repairXone participant's signal that another's message was infelicitous, incompletely understood, or otherwise problematic and for third-position repairXthe message-creator's attempt to redress miscommunication after her interlocutor's response reveals that she was misunderstood. Both types of repair have been the subject of a large body of literature on spoken language interaction. Little research has considered these types of repairs in the visual-gestural modality, though exceptions include studies on signed languages and on repair gestures that may accompany speech. The gestures observed in these studies have not been compared. We address this gap by considering gestural forms of repair initiation in an additional signed language, and comparing those observed in this language to those described in the prior literature.

This presentation offers a preliminary account of politeness in Ban Khor Sign Language, with particular attention devoted to analyzing how gestures contribute(d) to expression of politeness in the language. Beginning with examination of a highly conventionalized and ubiquitous gesture in Thailand, this presentation demonstrates how the wai and other emblematic gestures have become lexicalized in BKSL, in the process shedding all obligatory co-speech requirements yet still mirroring the original semantic co-speech associations. Additionally, by attending to Thai cultural norms for respectful physical comportment, this analysis illustrates how some norms (e.g., speaker's head elevation relative to the status of the interlocutor or avoidance of foot pointing) are maintained and manifested as politeness in BKSL, while other norms (e.g., modest eye gaze and equanimous facial expression) are at odds with basic grammatical requirements of language expressed in the manual modality.

We find that the unskilled hearing signer produces open repair initiators (i.e., those that do not specify the source of the trouble, and simply identify the prior talk as problematic]) through use of head tilt and brow furrowing. Notably, the head tilt we observe is formally similar to a second-position repair gesture that has been observed in non-signing spoken language users). For third-position repair, we observe deaf and hearing signers of all skill levels frequently beginning with one of several gestural emblems signifying negation and used alongside or in place of speech, and used as lexicalized negators in CSL, followed by repetition of the message the signer believes was misunderstood by the interlocutor. This is consistent with findings for spoken language repairs in this

position.

We discuss implications of our findings for future research on repair in the visual- gestural modality, giving particular attention to those gestural repair initiators that are common to repairs performed in CSL, in other documented signed languages, and in gesture.

"Wai" gesture matters: understanding politeness in Ban Khor Sign Language

Angela Nonaka

Historically, "gesture" and "sign language" were synonymous, even for linguists who deemed the latter to be "merely derivatives of language"(Bloomfield 1933:142). That view changed circa 1960 with the establishment of Sign Language Linguistics, when new emphasis was placed on strictly delineating gesticulation from language expressed in the manual modality, a trend that predominated for many years. In more recent decades, scholars revisited the complex relationship(s) between gesture and signed language, primarily in relation to lexicalization in new sign languages. Now, as the papers in this panel demonstrate, new research is underway to investigate young sign languages interactionally, vis-'o-vis conversation and pragmatics. In these domains too, gesture is interwoven in intricate and nuanced ways with signed language emergence and use.

A case in point is Ban Khor Sign Language (BKSL), a village sign language in Thailand. Less than 100 years old, the language is a relatively young but full-fledged sign language. It exhibits the traditional subsystems of language: phonology, morphology, lexicon, syntax and semantics as well as pragmaticsXincluding politeness, a language universal (Brown & Levinson 1987) that has been extensively examined in spoken languages (Watts et al. 1992, Watts 2003, Davies et al. 2011) but largely ignored in sign languages (Pietrosemoli 2001, Quinn 2004, Hoza 2007, George 2011).

This presentation offers a preliminary account of politeness in Ban Khor Sign Language, with particular attention devoted to analyzing how gestures contribute(d) to expression of politeness in the language. Beginning with examination of a highly conventionalized and ubiquitous gesture in Thailand, this presentation demonstrates how the wai and other emblematic gestures have become lexicalized in BKSL, in the process shedding all obligatory co-speech requirements yet still mirroring the original semantic co-speech associations. Additionally, by attending to Thai cultural norms for respectful physical comportment, this analysis illustrates how some norms (e.g., speaker's head elevation relative to the status of the interlocutor or avoidance of foot pointing) are maintained and manifested as politeness in BKSL, while other norms (e.g., modest eye gaze and equanimous facial expression) are at odds with basic grammatical requirements of language expressed in the manual modality.

Finally, BKSL originally developed and traditionally was used in particular sociolinguistic circumstances. i.e., by fellow Deaf and hearing Ban Khorians; independent of the national sign language, Thai Sign Language; and uninfluenced by formal deaf education or professional interpreting services. A language isolate, BKSL is distinct from all other languages in the community. Yet remarkable similarities remain in (meta)pragmatic awareness of politeness norms between BKSL signers and speakers of other languages in the village and elsewhere in Thailand. Thus, the Ban Khor case study provides new insights into the seamless ways in which shared gestures, especially those widely seen and used from birth by both hearing and deaf people inform language in the manual-visual modality, albeit in different ways, historically linguistically in the life of a sign language as well as interactionally in conversation.

Lexicalization of negative gestures in Chatino Sign

Lynn Hou, Kate Mesh

Chatino Sign Language is a young language that emerged spontaneously from a constellation of isolated home sign languages among 10 deaf people in two neighboring Chatino villages in rural Oaxaca. CSL has a stable lexicon of 6 manual negative particles, that can occur with and without accompanyinng non-manual signals. Three of the negative particles are used for basic clausal negation to make denials, rejections, and contradictions. While there is no clear one-to-one mapping between form and function, individual deaf signers exhibit strong preferences for one or two negators for basic clausal negation (Hou & Mesh, 2013a).

Three negators have been documented to occur across unrelated sign languages and linked to manual gestures of surrounding speech communities (Zeshan, 2004) while two negators are analyzed as emblems that are widely known across different countries (Matsumoto & Hwang, 2013). Only one negator appears to be a regional emblem in Oaxaca. All those negators have been used as conventionalized gestures in various communicative contexts, including signna 'o'ove monolingual Chatino speakers who communicated with Spanish speakers and different indigenous Mesoamerican speakers for negotiating trade in urban areas.

A micro-analysis of a 90-minute corpus of manual negative expressions in naturally-occurring conversa- tions between 5 deaf Chatino signers and 4 hearing co-signers of varying proficiency reveal three differences in how groups produce clausal negation (Hou & Mesh, 2013b). First, deaf signers tend to use negators more for clausal negation than negative interjection whereas hearing signers tends to produce negators in isolation. Second, deaf signers use multiple distinct negators as a possible strategy for forming negative intensives. Third, deaf signers tend to repeat one negator in one clause more than hearing signers do. We attribute some observed differences between deaf and hearing signers to the influence of clausal negation in spoken Chatino, the first language of most hearing signers.

What deaf and hearing signers share for expressing nega-

tion is that both groups frequently produce the negator in post-predicate and clause-final position. The influence of spoken Chatino cannot account for this, as negation tends to occur at the beginning of the clause and always precede the predicate in spoken Chatino. This raises empirical questions about how negative signs are produced as conventionalized gestures, namely as emblems, by sign-na 'o'ove hearing people. All six negators have been observed as conventionalized gestures, supporting the idea that gesture plays a crucial role in the development of sign languages (Wilcox, 2004).

Yet the process of how emblems are lexicalized in an emerging sign language, particularly with respect to the syntactical positioning of emblems in clauses, is not well-understood. Through metalinguistic interviews, we investigate how sign-na 'o'ove hearing Chatino and Spanish speakers interpret the meaning and function of negative emblems and how they position the emblems with accompanying speech. Understanding how sign-na 'o'ove hearing people produce negative emblems can link us better to understanding how emblems are lexicalized in an emerging sign language.

Brain 2: Talks

Gestures and communicative intent amongst autistic individuals

Rachel Chen, K.K. Luke

Impairment in social interaction is generally regarded as a key criterion for the diagnosis of Autism (American Psychiatric Association, 2013). Studies of Autistic individuals have noted their repetitive or stereotypical use of gestures as an indicator of their social impairment. (Mundy et al. 1986, Loveland et al. 1986, Ozuyurek at al. 2007, Silverman et al. 2010) Others examine their ability to imitate gestures (Aldridge et al., 2000; Ohta, 1987; Ingersoll, 2005; Roeyeus, 1998), but characterize it as idiosyncractic or inappropriate (Ham et al., 2008; Buffington et al., 1998 etc.). Still other studies have questioned autistic individuals' very possession of "communicative intent"(Kanner, 1943; Churchill, 1972; Rutter, 1984).

However, most data used in these studies were obtained in structured or quasi-structured sessions in an unfamiliar environment. Research has illustrated the importance of having a natural environment (Brown et al., 2008) and familiar interlocutors (Theodorou et al., 2010) for the facilitation of spontaneous interactions of autistic individuals. Furthermore, there has been little understanding of the sequential location of their gestures within interaction, which limits our understanding of how their motor actions can be communicative. Close examination of autistic individuals' interaction within a natural setting can therefore further our understanding of the way gestures may contribute to the performance of communicative actions.

In the present study, we analyse in some detail the gesturing of autistic individuals in interaction with others, with an aim to assessing the extent to which the gestures they produce are oriented to by their co-participants as meaningful. Under investigation are several individuals in Singapore, each diagnosed with Autism, in a range of everyday activities at home and in other familiar environments. Video recordings of naturally occurring interactions were obtained and transcribed following Conversation Analysis (CA) conventions, which provides an analytical framework that allows for a deeper and more nuanced understanding of such encounters (Dobbinson, 2010).

The study found that some aspects of the individuals' nonverbal behaviour which might otherwise be deemed idiosyncratic or self-stimulatory ("stimming") turned out on closer scrutiny to play a significant role in interaction, e.g., as a delaying device or to change the direction of an interactional sequence. Head movements, eye gazes and changes in body posture were also used selectively to meet communicative goals.

While many produced gestures involved directly manipulating another's hand, (similar to findings by Stone et al., 1997), autistic individuals also produced a significant amount of spontaneous gestures such as pointing, or iconic gestures in reference to requested objects. During a game situation, a nonverbal autistic individual not only jointly collaborated with his interlocutor by imitating, but also initiated new moves of his own, accompanied by smiling and shifts in eye gaze. These displays demonstrate their ability to comprehend their interlocutors' communicative moves, and to use gestures to pursue communicative goals within interaction.

When considering where these actions occur within a sequential interaction, the gestures produced display their engagement with their interlocutors. Their gestures jointly build upon shared sequences, and are meaningful productions within their relevant contexts. These and other examples suggest that one may attribute, with justification, more communicative intent to autistic individuals than has previously been acknowledged.

Distinguishing gesture processing from sign language processing: The contributions of the superior temporal lobe

David Corina, Laurie Lawyer, Michelle Cohn, Shane Blau

Studies of spoken language processing have observed exquisite differentiation of speech from non- speech sounds in the bilateral posterior superior temporal lobes. These anatomical regions have been claimed to show the initial sites of linguistic specialization for acoustic processing (Hickok & Poeppel 2007). In this study we examined whether similar temporal lobe regions were modulated by linguistic and non-linguistic gestures in deaf signers and hearing non-signers. We used fMRI techniques to explore the differences in the processing of individual signs and non-linguistic self-

grooming gestures in a novel implicit gesture recognition task. Our goal was to ascertain the regions in the temporal lobes that minimally differentiated these two forms of gesture. We focus our analysis on the superior temporal gyrus (STG) as this region has been routinely observed to differentiate speech from otherwise complex acoustic non-speech stimuli (Vouloumanos et al 2001; Narain, et al 2003). Subjects included 18 deaf signers (14 native, 4 early-signers) and 18 hearing non-signers, who during fMRI scanning (3T Trio Siemens, 3.63mm, TR 3000 ms.,TE 30 ms.) monitored video-clips of ASL signs and self-grooming gestures. For each exemplar, they indicated whether one or two hands were active during the actions by means of a counterbalanced key press. This paradigm allows each group to attend and respond to all stimuli, even though the hearing subjects were sign-naive. As the task does not require overt identification of the stimuli and does not require in depth linguistic analysis it is useful for identifying regions that play a fundamental role in the first steps of linguistic gesture processing. Our group level comparisons ($p < .01$) showed hearing subjects relative to deaf signers did not show areas of activation in the STG that differentiated ASL signs from gestures (each contrast relative to fixation). In contrast, deaf subjects relative to hearing subjects showed activation in the left hemisphere STG for signs (MNI coordinates: -67, -36, 8), and bilateral STG for gestures (-53, -36, 8; 66, -40, 11) (each contrast relative to fixation). Closer examination of the individual group data reveals that the deaf subjects showed bilateral STG activation for sign language (-67, -36, 8; 59, -36, 0) with activations for gestures that lie nearly adjacent but typically more posterior (-56, -44, 4; 66, -44, 8) to sign language peaks (all p's $<.001$). These bilateral regions lay well within published regions of interest that differentiate speech from non-speech signals (e.g. -58 (+/-4.8), -33(+/-10.3), -10(+/- 4.1) (Vouloumanos et al 2001, Narin et al 2003). Taken together these data provide evidence for a linguistic specification of the posterior superior temporal lobe that is agnostic to language modality. In addition, the lack of complementary STG activations in the hearing non-signers for signs or gestures further suggests this region is not a latent "gesture"area that has become specialized for linguistic processing.

Autism spectrum disorder and gesture production: Correlations between severity and interactive gesture

Inge-Marie Eigsti

Gestures serve a variety of distinct functions, facilitating 1) a listener's comprehension of a speaker's discourse, and 2) a speaker's own discourse formulation. A number of studies have revealed significant individual differences in gesture production that map onto specific cognitive and sociocommunicative characteristics (e.g., Chu, Meyer, Foulkes, & Kita, 2013). This correlational research has indicated that individuals with low visual working memory, spatial transformation and conceptualization abilities tend to produce more representational (iconic, metaphoric, and deictic) gestures. In contrast, individuals who are high in empathy tend to produce greater numbers of "conduit"and "palm-revealing"gestures, which serve an interactive function.

The current study uses a new approach to examine relationships between gesture and these important individual differences in cognitive and sociocommunicative characteristics. Specifically, individuals with autism spectrum disorder (ASD) and controls (n = 15 per group) matched on chronological age and full-scale IQ completed a battery of gesture-elicitation and subject characterization tasks.ASD is a neurodevelopmental disorder with particular relevance for the present study, because individuals with ASD have significant sociocognitive deficits, including impairments in representing the mental states of other individuals (Baron-Cohen, 1988), from a very early point in development (Mundy, Sigman, & Kasari, 1990). As such, study of affected individuals affords the opportunity to examine notions of causality, such as whether impairments in social cognitive are causally implicated in reduced rates of interactive gesture; and to examine relationships between non-literal language abilities and representational gestures.

Participants completed a definitions task, designed to elicit high rates of gesturing, in which they had to define six non-literal phrases ("get back in the saddle").Performance was coded for gesture type: interactive (conduit and palm-flip gestures) and representational (iconic, metaphoric, and deictic gestures) gestures. With analyses completed for 10 subjects to date, results showed no difference in total number of gestures, $p = .03$, consistent with prior research (de Marchena & Eigsti, 2010).The ASD group was less accurate in their definitions, 31% versus 50% correct in the control group, $p = .05$, though in this preliminary dataset the contrast did not reach significance.Finally, while representational gestures were produced at similar rates across groups, audience-oriented or interaction gestures differed as a function of group, and furthermore, were specifically correlated with ASD symptom severity, $r(10) = 0.64$, $p = 0.04$, despite similar performance on the task itself.Further analyses will examine whether accuracy on specific definitions was associated with gesture production during those trials, and how gestures relate to standardized language assessment scores.

Examination of the relationships between sociocognitive characteristics and gesture production illuminate the role that gesture plays in these characteristics. We found that individuals with ASD gesture as frequently as their peers, overall, but that their production of gestures with a specifically social function is associated with diagnostic severity.In addition to clarifying the specific nature of the gestural profile in ASD, a subject requiring significantly greater clarity, this approach and these results shed light on individual differences in gesture production, suggesting that sociocommunicative characteristics likely lead to differences in gesturing.

Superior temporal sulcus connectivity and the processing of speech and gesture

Benjamin Straube, Tilo Kircher

The left superior temporal sulcus (STS) plays an important role in integrating audiovisual information and is functionally connected to disparate regions of the brain. For the integration of gesture information in an abstract sentence context (metaphoric gestures, MP) connectivity between the left STS and the inferior frontal gyrus (IFG) should be important. By contrast reduced connectivity can be expected for the integration of gesture information in a concrete sentence context (iconic gestures, IC). Thus, we tested the hypothesis that the functional connectivity of the left STS is dependent on modality (speech vs. gesture) and abstractness of speech-gesture utterances (IC vs. MP).

During fMRI-data acquisition, 16 participants were shown videos of an actor performing gestures in a concrete (IC) and abstract (MP) sentence context. Additionally unimodal speech (S) and gesture (G) conditions were presented. A psycho-physiological interaction (PPI) analysis based on the seed region from a previous analysis in the left STS was performed (see Straube, et al., 2013).

Across all conditions (IC, MP, S, G) we found common positive connectivity of the STS seed region to the left posterior temporal cortex, the right parietal cortex and the bilateral inferior frontal gyrus. Modality specific STS connectivity were found for speech processing (S¿G) in language related areas and gesture processing (G¿S) in occipital areas. For MP¿IC we obtained increased connectivity in predominantly lateral and medial frontal brain regions. By contrast, for IC¿MP we found increased connectivity in occipital brain regions. Finally reduced connectivity for iconic coverbal gestures in contrast to both unimodal conditions (IC IC) were found in bilateral and medial frontal brain regions.

These data suggest that the STS is generally connected to the bilateral IFG and the left posterior temporal cortex, supporting the assumption that this region is especially relevant of for the semantic processing of speech and gesture information. Modality specific effects indicate that the STS receives information from the visual (G condition) and auditory system (S condition), which supports the suitability of this region as a potential integration site. Finally, the reduced connectivity in the IC condition suggests a crossmodal facilitation effect. Thus, frontal brain regions are less relevant when congruent speech and gesture information are presented together. This is especially true for gestures in a concrete sentence context. In addition to the often reported relevance of the left IFG for speech gesture integration (Kircher, et al., 2009; Straube, et al., 2011; Straube, et al., 2009; Willems, et al., 2007; Willems, et al., 2009), the connectivity data suggest that the right IFG also contributes to STS function - especially for the unimodal processing of speech and gesture or the processing of gestures in an abstract sentence context (MP).

Metaphor 2: Talks

The past and future are in your hands: How gestures affect our understanding of temporal concepts.

Melvin Ng, Winston Goh, Melvin Yap, Chi Shing Tse, Wing Chee So

It is not uncommon to use metaphors to represent and reason about time in our daily conversations. In English, metaphors referring to time are arranged along the sagittal axis (e.g., N'o'oez & the future lies ahead of us, we often look back on our past; Clark, 1973;). In addition to speech, speakers also gesture about time while speaking. Interestingly, however, Cassanto and Jasmin (2012) have found that speakers tended to gesture predominantly along the lateral axis in spontaneous speech (e.g., pointing to the left to represent the past and to the right to represent the future). Sagittal gestures (e.g., pointing to the front for the future and to the back for the past) were observed as well, though in much smaller proportions during spontaneous speech and more prominently during deliberate descriptions. Boroditsky (2001) and Chui (2011) presented another possible axis along which our concept of time may be aligned as well in Chinese speakers: the vertical axis. The question of interest in this study is whether the conceptual representation of time is predominantly aligned against sagittal, lateral, or even the vertical plane among English- speaking adults. Previous studies have shown that gestures prime semantically related words and concepts (Yap et al., 2011; Wu & Coulson, 2011), thus suggesting that gesture allows us to investigate implicit spatial conceptualization of time. We aim to examine whether the presentation of a pointing gesture (e.g., pointing to the left) would facilitate classification of a semantically related temporal word (e.g., yesterday), using a cross-modal semantic priming paradigm. In the present study, we asked adult speakers to watch video clips where the model produced pointing gestures (e.g., pointing to the left) while narrating temporal words which were either past-related (e.g., yesterday) or future-related (e.g., ahead). Each video clip lasted for approximately 3500 milliseconds. The pointing gesture and its accompanying auditorily presented temporal words were temporally synchronized. Participants were randomly assigned to three conditions which differed in the direction of pointing gestures. In the lateral condition, the model pointed to her left and right; in the sagittal condition, the model pointed Sweetser, 2006 to her front and back; in the vertical condition, the model pointed up and down. Participants in all three conditions were required to determine whether the auditory tokens were related to the past or future. We measured their response accuracy and reaction time. It is expected that the presentation of congruent pairs of pointing gestures and their co-occurring temporal words (e.g., point to the left and "past") would result in faster reaction time than the presentation of incongruent pairs (e.g., point to the

right and "past"). Results obtained from 15 participants in a recent pilot reveal effects of congruency for future-related words but not past-related words when primes are temporal gestures presented along the lateral axis. Further investigations are underway to determine why there is an absence of congruency effects for past-related words.

Gesturing the source domain: Exploring the metaphorical models of transgenderism

Jenny Lederer

Gesture is aptly described as a "backdoor"to cognition (Sweetser 2007: 203). Co-speech gesture has been shown to encode metaphorical source domains (Cienke 1998), aid in the representation of abstract concepts (Perril and Sweetser 2004), and specific handshapes, movements, and directionality systematically structure metaphorical vocabulary in American Sign Language (Taub 2001). Although gesture is a rich source of data for the examination of conceptual metaphor, it is noticeably absent from the critical and political discourse analysis paradigms. In this presentation, I use gesture to investigate which source domains are structuring American understandings of transgenderism, the concept ascribed to those who have begun or completed a change in their sex characteristics from male to female or female to male. Through the examination of twenty transition narratives documented on video, I will show how both co-speech gesture and an emerging lexicon of ASL signs align with spoken and written narrative to support a spatially based representation of gender identity and transition. Recently, there has been a large amount of work analyzing the construction of transgender identity (e.g. Armitage 2008; Valentine 2007), some of which includes linguistic analyses of transgender, transsexual, and drag queen communicative patters (Barrett 1998, 1999). However, there exists no comprehensive analysis of the cognitive models used to understand transgender identity or the transition process. I offer a roadmap for those interested in incorporating evidence from gesture into the identification of unconscious assumptions, which organize speakers' comprehension of complex political topics.

The assignment of gender is talked about and thought about as being located in a bounded region; English speakers qualify and quantify gender and transition through their understanding of movement through space: cross-dressing, transitioning, changing, male-to-female, coming out, intersex. This language is indicative of a binary category model of gender assignment, in which each category is understood as a bounded region in space and transition is a journey with intermediate and final destinations along a path as in (1):

(1) I have often likened my transition to slowly wading out into a cold lake. I take a step or two, shiver a bit at the coldness, and hang out for a bit as my body acclimates. Then I decide if I want to go deeper. All along the transition I have been open to the concept that I can stay where I am, go back, or push deeper. And though several times I have pulled back too deep, too fast, too cold- I have always found myself moving toward transition.

Co-speech gesturing from my corpus, such as two upward facing palms in alternate motion, canonical of decision-making (MAKING DECISION IS WEIGHING), suggests the coming out process is understood as a choice with two alternatives. In one specific example of this gesture, the two palms are subsequently coopted into deictic reference points on the left to right timeline. The temporal reference set up in the gesture signals a spatial threshold, which once passed, cannot be re-traveled.

Does physical co-presence mediate the effects of gesture on spatiotemporal metaphor?

Tasha Lewis, Elise Stickles

Gesture serves an important communicative purpose in human interaction, including metaphor priming, as metaphoric gestures motivate spatiotemporal metaphor use (Stickles and Lewis 2013). In face-to-face communication (FFC), gestures toward the speaker prime TIME IS MOTION PAST EGO; the addressee (taking the speaker's perspective) thinks of time as moving towards the self. Previous work shows that visibility is also important in FFC: the use of communicative gesture can be affected by the lack of visibility (Bavelas and Chovil 2000). Work on computer-mediated communication (CMC) shows that video-based communication reduces gesture rate and size (Mol, Krahmer, Maes, and Swerts 2011). Here we consider the effects of CMC on gestural priming: does a lack of physical co-presence alter attention to metaphoric gesture.

Participants were asked an ambiguous question, eliciting different responses ("Monday"or "Friday") depending on the participant's spatiotemporal metaphor in use. This was paired with gestures which are either congruent from the speaker's perspective with the ego-motion (TIME IS EGO MOVING THROUGH SPACE) variant (gesture away from speaker), the temporal-motion variant (toward the speaker), or neither (lateral, reflecting the typical gesture English speakers make accompanying temporal language (Casasanto 2012)). 374 participants on Mechanical Turk (www.MTurk.com) watched a video of a speaker asking the question while gesturing in one of several conditions: no gesture; gesture away from the speaker, palm upright or sideways; toward the speaker, palm upright or sideways; from the speaker's right-to-left, palm upright or sideways; or from left-to-right, palm upright or sideways. 80 viewed from a third-person viewpoint and 294 viewed from first-person.

Results were analyzed using a contrast-coded multiple logistic regression model. Gesture direction (away/towards) significantly improved the fit of the model ($p < 0.05$). The "away"gesture is significantly more likely to elicit a "Monday"response (odds ratio = 2.521, $p < 0.001$). There were no effects of viewpoint, palm orientation, or right/left gesture (p ¿ 0.05). Viewpoint, palm orientation, and lateral gestures do not convey information regarding forward motion. There-

fore those conditions did not influence participants to use one metaphor over another, as shown by those factors' lack of significant effects. Gestural information that is relevant to forward motion i.e. the towards and away gestures did influence addressees' metaphor use.

Stickles and Lewis 2013, which performed the same task in FFC, found addressees used metaphors congruent with motion from the speaker's perspective. Here participants responded using metaphors congruent with the addressee's perspective: a gesture away from the speaker is for the addressee congruent with the temporal-motion metaphor, which accords with the higher rate of "Monday"responses. In the case of CMC, a lack of a shared physical discourse space influences the addressee's ability to interpret the information conveyed by gestures which are dependent on relative frames of reference. While addressees in FFC can reliably take the perspective of the speaker with regards to forward-backward motion, in CMC they instead maintain that of the addressee, presumably due to the lack of shared physical ground caused by the lack of physical co-presence.

Mixing metaphors in co-speech gesture

Esther Walker, Kensy Cooperrider

What insights can gesture offer into the nature of metaphorical thought. Much metaphor research has focused on "orientational metaphors (Lakoff & Johnson, 1980) in which people recruit basic contrasts from the domain of space up and down, front and back, and left and right for thinking and talking about concepts in more abstract domains like number, valence, and time. Curiously, there is often more than one orientational metaphor available for a given abstract concept: numbers can be higher but also further to the right along the mental number line, good can be up but also associated with our dominant-hand side, and the future can lie ahead but also off to our right. These metaphors have different experiential origins and previous experimental designs have embodied the assumption that they are distinct. In this presentation, using orientational metaphors for time as a case study and gestural evidence as a window, we challenge this assumption.

Informal observations suggest that speakers sometimes gesture in a way that mixes front-back and left-right metaphors for time, for example when referring to the future with a gesture directed both forward and rightward or to the past by pointing backward over the left shoulder. We conducted two studies to investigate how pervasively and systematically English speakers mix these metaphors in their gestures. Study 1 used an explicit gesture elicitation paradigm in which participants (n=104) were directly asked how they would gesture about the past or future (Casasanto & Jasmin, 2012). Of the gestures produced along a clearly codable axis, 67% were produced along the front-back axis, 16% along the right-left axis, and 17% of gestures combined these axes. Crucially, combined-axis gestures were doubly congruent (forward-rightward for future; backward-leftward for past) 79 % of the time, strongly suggesting they are more than mere motor noise. Moreover, for congruent gestures that included a forward or backward component, participants' choice of the right or left hand was modulated by whether they were gesturing about the past or future (e.g., left-handed gestures were used more often for the past). Because the mixing of metaphors observed in Study 1 may be due to its peculiar task demands, Study 2 used a more implicit elicitation procedure. Participants (n=24) were asked to define concepts (8 target temporal concepts; 24 filler concepts) with no instructions about gesture. Temporal gestures were spontaneously produced along the front-back axis (27%), along the right-left axis (72%), and in ways that combined these axes (11%). Parallel to our observations in Study 1, combined-axis gestures were doubly congruent 77% of the time. Furthermore, for forward and backward gestures, participants' hand selection was again modulated by whether they were gesturing about the past or future.

Our results suggest that orientational metaphors are best thought of- X not as inviolable wholes- X but as something more like attractors in a continuous state space (Spivey, 2007). Moreover, the findings provide another demonstration that gestures, as fine-grained three-dimensional spatial structures that unfold in time, offer insights into mental representation that have no counterpart in spoken language.

Disfluency: Talks

When speech stops, gesture stops: evidence from crosslinguistic and developmental comparisons

Maria Graziano, Marianne Gullberg

Some theoretical proposals concerning the gesture-speech relationship suggest that gestures mainly have a compensatory function, that is they are employed to aid lexical retrieval (Krauss et al., 2000), conceptualisation, or information packaging (Alibali et al., 2000; Kita, 2000). This notion can also be found in many developmental studies of both first and second language learning. These assumptions predict that gestures should be more frequent during disfluent than fluent stretches of speech. However, little is known about the relationship between gestures and fluent vs. disfluent speech, and whether it varies across languages, or between competent and developing language users, whether children or adults. The aim of study is therefore to investigate the putative compensatory role of gestures by examining in close detail the gesture production in adult speakers of two different languages (Italian and Dutch) and in two types of language learners (child and adult second language learners). More specifically, we aim to explore the following research questions: (1) do speakers of different languages or linguistic competence predominantly produce gestures with fluent or with disfluent speech. (2) what articulatory features do gestures have during disfluent speech. (3) what functions do gestures completed dur-

ing disfluencies have. Analyses were conducted on narrative retellings produced in dyadic, interactive settings by 11 adult Italian and 11 adult Dutch native speakers; 33 Italian children in three age groups (4-5; 6-7; 8-10 years), each containing 11 subjects, and 11 Dutch adult learners of L2 French at low to intermediate levels of proficiency. All spoken disfluencies were identified defined as filled and unfilled pauses, interruptions, and lengthening. All gestures were identified and coded for whether they occurred with fluent or disfluent speech. Gestures were further coded for structural properties (complete vs. interrupted stroke), and for function (referential vs. pragmatic gestures). The results show no crosslinguistic and no developmental effects. Instead (1) all groups, child and adult learners and adult native speakers of Italian and Dutch alike, predominantly produce gestures during fluent speech and only rarely produce gestures during disfluencies. However, L2 learners are significantly more likely to do so than the other groups; (2) in all groups gestures during disfluencies tend to be suspended; (3) in all groups the small number of gestures completed in disfluencies are both lexically related (referential gestures) and metalinguistic comments on communication breakdowns (pragmatic gestures). Overall, the data strongly suggest that when speech stops, so does gesture in adult speakers of different languages as well as in developing language users, whether they are children or adults. The findings constitute an important challenge to both gesture and language acquisition theories assuming a mainly (lexical) compensatory role for (referential) gestures, and provide strong support for the notion that speech and gestures form an integrated system.

Coordination between the hands and mouth: A kinematic experiment in ASL

Jonathan Udoff, Ignatius Nip, Karen Emmorey

Signed languages offer a unique situation in which to study the precise coordination of articulatory gestures. Much of the coordination of independent articulators that occurs in spoken languages - speech with co-speech gestures- happens at the suprasegmental level or higher (Wagner, Malisz, & Kopp, 2014). In contrast, signed languages frequently employ "cosign" mouthings- unvocalized lip movements that correspond to a spoken language translation of the manual sign- that are produced contemporaneously with the sign source (Bank, Crasborn, & van Hout, 2013; Boyes Braem & Sutton-Spence, 2001). While there may be a communicative pressure to produce the two articulations in a simultaneous fashion, there may also be neuromotor constraints on how certain kinds of movements may coordinate with others across articulatory channels (e.g., the hands and mouth; Woll, 2001) Particularly, the cortical proximity of hand and mouth motor areas and their connections to Broca's area may cause movements of the hands to influence simultaneous movements of the mouth during language production (Gentilucci, Benuzzi, Gangitano, & Grimaldi, 2001; Gentilucci & Volta, 2007).

The present study seeks to understand the factors that contribute to the fine-grained coordination of manual gestures with oral gestures during the production of American Sign Language. The experimental task asked 10 native deaf signers to simultaneously fingerspell and mouth nonce words, designed to elicit a variety of mouth-hand movement combinations. Movements of the hands and mouth were recorded using a passive marker optical motion capture system. The system tracked each of ten markers affixed to participants' hands and face to calculate their position in three-dimensional space. Movement signals were created that reflect the opening and closing of the dominant (right) hand and of the lips. A cross-correlation analysis of the contemporaneous movement signals provide two measures of coordination between the two articulators: correlation coefficient (r) values close to zero indicate a low degree of spatial coupling while high lag values indicate a low degree of temporal coupling.

Overall, movements of the two articulators are tightly coupled in space and time: r=0.79, lag = 74 ms; however, further analysis revealed a number of effects. First, mouthings exhibit greater coordination with the manual fingerspelling when both the hands and mouth require the same number of gestures. This effect is observed for both the spatial (r=0.83 vs. r=0.73, $t(9)=6.5$, $p=0.0001$) and temporal measures (57 ms vs. 106 ms lag, $t(9)=6.4$, $p=0.0001$). Second, mouth-hand movement combinations of one gesture are more tightly coordinated than combinations of two gestures. This effect applies to both spatial and temporal coordination: r=0.86 vs. r=0.78, $t(9)=7.3$, $p<0.0001$; 38 ms vs. 87 ms lag, $t(9)=5.2$, $p<0.001$. Contrary to previous accounts from co-sign mouth gestures (Woll, 2001), there is no effect of the congruency of gesture direction between the hand and mouth. Items that require the hands and mouth to move in different directions were not significantly less coordinated in space or time than items that require the two body parts to move in the same manner: r=0.84 vs. r=0.82, $t(9)=2.2$, $p=0.052$; 60 ms vs. 55 ms lag, $t(9)<1$, $p=0.56$.

These findings suggest that successful coordination between the hands and mouth is dependent on the number of gestures required by each articulator, while the specific movement that defines the gestures is largely irrelevant. This work is an important step in expanding our understanding of singleton gesture production to be able to explain more complex concatenations involving multiple articulators.