

Provisional schedule 2020.03.04

# AttLis 2020

Storrs, CT, USA

March 17-18



## The Attentive Listener in the Visual World

### *Invited speakers*

**Elika Bergelson**  
*Duke University*

**Craig Chambers**  
*University of Toronto*

**Mike Tanenhaus**  
*University of Rochester*

**Mila Vulchanova**  
*Norwegian University of  
Science & Technology*

[AttLis](#) addresses multi-modal cognition with an emphasis on the interaction between language and vision. More specifically, there is a central focus on how attentional and visual processes interact with spoken and written language processing. Why are attention and vision crucial to language comprehension? How does each inform and mediate the other in moment-to-moment processing and in language development over the lifespan?

*This will be the seventh AttLis meeting, and the first in North America.*

*This meeting is supported by the Science of Learning & Art of Communication interdisciplinary Ph.D. training program at the University of Connecticut*

<https://slac.uconn.edu/attlis-2020>

**UConn**  
SCIENCE OF LEARNING  
& ART OF COMMUNICATION

## BRIEF SCHEDULE

TUESDAY, MARCH 17, UCONN STORRS, OAK HALL 101

9:00	<b>Breakfast</b>
9:45	Opening remarks
9:50	<b>Aspects of language divergently map onto motor deficits, dyspraxia and symptom severity in children with autism</b> Bahar Tuncgenc, Carolyn Koch, Mackenzie Stabile, Stewart Mostofsky & Inge-Marie Eigsti
10:15	<b>KEYNOTE</b> <b>Different ways of making a point: a study of gestural communication in typical and atypical early development</b> Mila Vulchanova
11:10	<b>Timing visual attention for language acquisition in interactions between deaf children and their parents</b> Amy Lieberman, Allison Fitch & Eric Setzer
11:35	<b>Break</b>
11:50	<b>Predictive processing of grammatical gender in bilingual children</b> Jasmijn Bosch & Francesca Foppolo
12:15	<b>How does bilingual reading guide object identification of visual referents?</b> Naomi Vingron, Madelaine Thomas & Debra Titone
12:40	<b>Lunch served outside venue</b>
14:10	<b>The moment-to-moment pitch dynamics of child-directed speech shape toddlers' attention and learning</b> Mira Nancheva, Elise Piazza & Casey Lew-Williams
14:35	<b>Pupil dilation reflects attentional effort in processing clear and conversational speech</b> Hannah Mechtenberg & Emily Myers
15:00	<b>Break</b>
15:20	<b>Online-processing of protagonists' perspective-taking</b> Sara Meuser, Stefan Hinterwimmer & Maximilian Hörl
15:55	<b>Effect of coordination on perspective taking: Evidence from eye-tracking</b> Yipu Wei & Yingjia Wan
16:10	<b>Empathy influences behavioral perceptions and eye movements in non-literal language processing</b> Pavitra Makarla, Lauren Benson, Hana Kim, Gitte Joergensen & Kathrin Rothermich
16:35	<b>Break</b>
16:50	<b>KEYNOTE</b> <b>The nascent lexicon: Comprehension in the lab and in the world</b> Elika Bergelson
17:55	<b>Break</b>
18:15	<b>Dinner served in Bousfield</b>

WEDNESDAY, MARCH 17, UCONN STORRS, OAK HALL 101

8:15	Breakfast
	<b>KEYNOTE</b>
9:00	Intention, attention, and real-time language processing in little listeners Craig Chambers
9:55	The timecourse of spoken word recognition across development Elizabeth Schoen Simmons & James S. Magnuson
10:20	Break
10:40	The role of early language exposure and modality on bottom-up visual attention Allison Fitch, Sudha Arunachalam & Amy Lieberman
11:05	Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing Falk Huettig & Ernesto Guerra
11:30	Lunch
	<b>KEYNOTE</b>
12:30	Perception, Action and Language with Virtual Reality: Low hanging fruit or deluded speculations? Michael Tanenhaus
13:20	Closing remarks
13:40	Bus leaves Nathan Hale for UMass-Amherst

## AUTHOR DETAILS

Sudha Arunachalam	<a href="mailto:sudha@nyu.edu">sudha@nyu.edu</a>	New York University	United States
Lauren Benson	<a href="mailto:lvbenson@iu.edu">lvbenson@iu.edu</a>	Indiana University Bloomington	United States
Elika Bergelson	<a href="mailto:elika.bergelson@duke.edu">elika.bergelson@duke.edu</a>	Duke University	United States
Jasmijn Bosch	<a href="mailto:jasmijn.bosch@unimib.it">jasmijn.bosch@unimib.it</a>	University of Milano-Bicocca	Italy
Craig Chambers	<a href="mailto:craig.chambers@utoronto.ca">craig.chambers@utoronto.ca</a>	University of Toronto-Mississauga	Canada
Inge-Marie Eigsti	<a href="mailto:inge-marie.eigsti@uconn.edu">inge-marie.eigsti@uconn.edu</a>	University of Connecticut	United States
Allison Fitch	<a href="mailto:afitch@bu.edu">afitch@bu.edu</a>	Boston University	United States
Francesca Foppolo	<a href="mailto:francesca.foppolo@unimib.it">francesca.foppolo@unimib.it</a>	University of Milano-Bicocca	Italy
Ernesto Guerra	<a href="mailto:ernesto.guerra@ciae.uchile.cl">ernesto.guerra@ciae.uchile.cl</a>	Center for Advanced Research in Education, Universidad de Chile	Chile
Stefan Hinterwimmer	<a href="mailto:hinterwimmer@uni-wuppertal.de">hinterwimmer@uni-wuppertal.de</a>	Bergische Universität Wuppertal	Germany
Falk Huettig	<a href="mailto:falk.huettig@mpi.nl">falk.huettig@mpi.nl</a>	Max Planck Institute for Psycholinguistics, Nijmegen	Netherlands
Maximilian Hörl	<a href="mailto:mhoerl1@uni-koeln.de">mhoerl1@uni-koeln.de</a>	University of Cologne	Germany
Gitte Joergensen	<a href="mailto:gitte.joergensen@uconn.edu">gitte.joergensen@uconn.edu</a>	University of Connecticut	United States
Hana Kim	<a href="mailto:kimha15@students.ecu.edu">kimha15@students.ecu.edu</a>	East Carolina University	United States
Carolyn Koch	<a href="mailto:kochc@kennedykrieger.org">kochc@kennedykrieger.org</a>	Center for Neurodevelopmental and Imaging Research, Kennedy Krieger Institute	United States
Casey Lew-Williams	<a href="mailto:caseylw@princeton.edu">caseylw@princeton.edu</a>	Princeton University	United States
Amy Lieberman	<a href="mailto:alieber@bu.edu">alieber@bu.edu</a>	Boston University	United States
James Magnuson	<a href="mailto:james.magnuson@uconn.edu">james.magnuson@uconn.edu</a>	University of Connecticut	United States
Pavitra Makarla	<a href="mailto:pavitra.makarla@uconn.edu">pavitra.makarla@uconn.edu</a>	United States	
Hannah Mechtenberg	<a href="mailto:hannah.mechtenberg@uconn.edu">hannah.mechtenberg@uconn.edu</a>	University of Connecticut	United States
Sara Meuser	<a href="mailto:smeuser2@uni-koeln.de">smeuser2@uni-koeln.de</a>	Universität zu Köln	Germany
Stewart Mostofsky	<a href="mailto:mostofsky@kennedykrieger.org">mostofsky@kennedykrieger.org</a>	Center for Autism and Related Disorders, Kennedy Krieger Institute	United States
Emily Myers	<a href="mailto:emily.myers@uconn.edu">emily.myers@uconn.edu</a>	University of Connecticut	United States
Mira Nencheva	<a href="mailto:nencheva@princeton.edu">nencheva@princeton.edu</a>	Princeton University	United States
Elise Piazza	<a href="mailto:epiazza@princeton.edu">epiazza@princeton.edu</a>	Princeton University	United States
Kathrin Rothermich	<a href="mailto:ROTHERMICHK17@ecu.edu">ROTHERMICHK17@ecu.edu</a>	East Carolina University	United States
Elizabeth Schoen Simmons	<a href="mailto:elizabeth.a.simmons@uconn.edu">elizabeth.a.simmons@uconn.edu</a>	University of Connecticut	United States
Eric Setzer	<a href="mailto:esetzer@bu.edu">esetzer@bu.edu</a>	Boston University	United States
Mackenzie Stabile	<a href="mailto:mackenzie.stabile@uconn.edu">mackenzie.stabile@uconn.edu</a>	University of Connecticut	United States
Michael Tanenhaus	<a href="mailto:mtan@bcs.rochester.edu">mtan@bcs.rochester.edu</a>	University of Rochester	
Madelaine Thomas	<a href="mailto:madelaine.thomas@mail.mcgill.ca">madelaine.thomas@mail.mcgill.ca</a>	McGill University	Canada
Debra Titone	<a href="mailto:debra.titone@mcgill.ca">debra.titone@mcgill.ca</a>	McGill University	Canada
Bahar Tuncgenc	<a href="mailto:Bahar.Tuncgenc@nottingham.ac.uk">Bahar.Tuncgenc@nottingham.ac.uk</a>	University of Nottingham	United Kingdom
Naomi Vingron	<a href="mailto:naomi.vingron@mail.mcgill.ca">naomi.vingron@mail.mcgill.ca</a>	McGill University	Canada
Mila Vulchanova	<a href="mailto:mila.vulchanova@ntnu.no">mila.vulchanova@ntnu.no</a>	Language Acquisition and Language Processing Lab, NTNU	Norway
Yingjia Wan	<a href="mailto:wanyj@psych.ac.cn">wanyj@psych.ac.cn</a>	Institute of Psychology, Chinese Academy of Sciences	China
Yipu Wei	<a href="mailto:yipu.wei@outlook.com">yipu.wei@outlook.com</a>	Peking University	China

## DETAILED SCHEDULE & ABSTRACTS

TUESDAY, MARCH 17, UCONN STORRS, OAK HALL 101	
9:00	Breakfast
9:45	Opening remarks
9:50	<p style="text-align: center;"><b>Aspects of language divergently map onto motor deficits, dyspraxia and symptom severity in children with autism</b></p> <p style="text-align: center;">Bahar Tuncgenc, Carolyn Koch, Mackenzie Stabile, Stewart Mostofsky &amp; Inge-Marie Eigsti</p>
	<p><b>Background.</b> Language deficits commonly observed in the autism spectrum conditions (ASC) span a wide range, including impairments in pragmatics (i.e., social-contextual aspects of language), word choice differences, and delays in production of grammatical structures. In storytelling, a critical skill for academic progress and social interactions, narratives of children with ASC have a shorter and less coherent structure as compared to those of typically-developing (TD) children (see meta-analysis by Baixauli et al., 2016). Children with ASC also use fewer mental-state terms (e.g., “know”, “think”) to express others’ perspectives, and have more difficulty comprehending social-emotional words (Tager-Flusberg et al., 2006). Certain discourse tools, such as first-person quoted speech, may also reflect perspective-taking abilities (Brunye et al., 2009), but whether first -person speech use is impaired in ASC requires further investigation. A recent longitudinal study (Bal et al., 2019) reported that early motor skills strongly predicted language skills at age 19, indicating the importance of examining relations between language and motor skills in ASC.</p> <p><b>Objectives.</b> This study examined (a) whether narrative structure and first-person speech usage is impaired in children with ASC, and (b) how linguistic variables correlate with autism-specific impairments: autistic trait severity (Social Responsiveness Scale-2: SRS-2), dyspraxia (child-modified version of Florida Apraxia Battery: FAB- child), and impaired basic motor control (Physical and Neurological Examination of Subtle Signs: PANESS).</p> <p><b>Methods.</b> Fifty-five 8- to 12-year-old children (29 ASD, 26 TD) watched and listened as a narrator (a trained actor) told a story; at periodic pauses in the narration, they retold the story. Key story elements were identified a priori; the proportion of key elements included in retellings was the measure of narrative structure. First-person speech relative to total utterances was calculated.</p> <p><b>Results.</b> Narrative structure and first-person speech were significantly correlated, <math>r = .57</math>, <math>p &lt; .0001</math>. Further, compared to peers with TD, children with ASC produced significantly shorter narrations, <math>F(2,52) = 10.358</math>, <math>p = .02</math>, and less first-person speech, <math>\chi^2(55) = 5.45</math>, <math>p = 0.02</math>. Impaired narrative structure was associated with increased ASD severity (SRS-2), <math>r = -.33</math>, <math>p = .02</math>, and poorer motor ability (PANESS), <math>r = -.49</math>, <math>p = .001</math>, but not with praxis. Decreased first-person speech production was associated with poorer praxis (FAB-child), <math>r = -.37</math>, <math>p = .05</math>, but not with motor ability or autistic trait severity.</p> <p><b>Conclusions.</b> Children with ASC showed reduced command of two language variables, narrative structure and first-person speech usage, during story retellings. Importantly, the two variables showed divergent patterns in their relationship with other autism-associated impairments. Strong associations of narrative structure with motor ability and ASC severity support previous research that narrative structure is a core social-communicative impairment in ASC. First-person speech being uniquely associated with dyspraxia may reflect the perspective-taking demand that is common in both measures. These findings provide insight into how different features of pragmatic language map onto the heterogeneous autism phenotype.</p>

## **KEYNOTE**

Tuesday  
10:15

### **Different ways of making a point: a study of gestural communication in typical and atypical early development**

Mila Vulchanova

Research over the past three decades has accumulated evidence suggesting that language and gesture are part of an integrated common system of communication (McNeill, 1992, 2015). For example, brain imaging studies in adults demonstrate that the process of meaning integration between speech and co-occurring gesture involves classic language areas in the left frontal and temporal lobes and their right hemisphere homologues (Andric & Small, 2012; Dick, Mok, Beharelle, Goldin-Meadow & Small, 2014). Gesture has been shown to play an important role in communicative development. Gesture development, and the production of deictic gesture pointing, in particular, both predates and predicts later language development in typically developing children (Iverson & Goldin-Meadow, 2005).

Deictic pointing gestures are the most severely impaired gestures in autism. Studies on typical development have proven that the shape of the hand and whether the referent is touched or not while pointing, are associated with different communicative intentions and developmental level. Despite the importance of these formal features, only one study exploring gesture types in toddlers at high risk for autism, included a subset of formal characteristics in their analyses (LeBarton & Iverson, 2016). The aim of the present study is to provide an exhaustive formal description of pointing gesture production patterns in autism with an improved pointing-gesture tailored methodology, attending to handshape (index/open palm) and contact with the referent (contact/no contact). Participants (age range=1-6 years old) with ASD (n=16), at high risk for autism (n=13) and typically developing children (n=18) were videotaped while performing a gesture elicitation interactive task with their caregivers. Results showed significant group differences between the autism and the typically developing groups in the total number of pointing gestures and in both formal features, handshape and contact with the referent. Children with autism produced fewer no contact index finger pointing compared to the typically developing children. These group differences are not attributable to parental input, as no statistically significant differences were found in caregivers' gesture production. These results confirm that there are not only quantitative, but also qualitative differences in gesture production by children on the autism spectrum. This has the potential to categorize autism-markers and add to the behavioural characterization of the condition.

Tuesday  
11:10

## Timing visual attention for language acquisition in interactions between deaf children and their parents

Amy Lieberman, Allison Fitch & Eric Setzer

Parental input in the early years is a robust and reliable predictor of vocabulary and other academic abilities (Rowe, 2012). Input that occurs during moments of shared joint attention enables children to make immediate connections between linguistic input and the surrounding visual world (Tomasello and Farrar, 1986), particularly when parents use a “follow-in” approach of labeling objects to which the child is attending (Akhtar et al., 1991). For hearing children, the association between language and its referents is made through simultaneous and multi-modal perception; language input is perceived through the auditory mode, and objects are perceived through the visual mode, the tactile mode, or both. In contrast, deaf children acquiring American Sign Language (ASL) perceive both linguistic and non-linguistic information through the visual mode. The social dynamics of gaze, attention-getting, and object labelling in these visually-complex interactions likely differ from those of spoken language, but have not yet been studied extensively (Lieberman et al., 2014).

In the current study, we analyzed naturalistic ASL-based interactions between deaf children ( $n = 18$ ) ages 9-60 months ( $M = 34$  months) and their deaf ( $n=12$ ) or hearing ( $n=6$ ) parents. We hypothesized that, if parents use a follow-in approach to joint attention, they would label objects to which children were previously attending. We further predicted that spontaneous gaze shifts to the parent would be positively correlated with the child’s age. Twelve minutes from each video were coded off-line for child-directed ASL signs, parent attention-getting cues, and the location of child gaze and object touch surrounding each ASL sign produced by the parent. We analyzed all concrete nouns ( $n = 791$ ) to determine how attention-getting, child gaze, and child object touch aligned with parent labelling. We first coded whether the noun was accessible, defined as whether or not the child was visually attending to the parent when the noun was produced. Next, for accessible nouns, we determined whether the child gazed at and/or touched the referent object within two seconds prior to the label, and whether the parent overtly elicited the child’s attention prior to the label.

Mutual gaze was established between the parent and child for 47% of concrete nouns, and proportion of accessible nouns was positively related to age ( $r = .50$ ,  $p = .034$ ). Parents used a follow-in approach by labelling an object to which the child had previously been attending (looking at or touching) 65% of the time, and this was positively correlated with age ( $r = .52$ ,  $p = .028$ ). Parental attention-getting cues were relatively rare, preceding 32% of object labels. This suggests that deaf children are adept at managing their own visual attention through spontaneous gaze shifts to the parent. Taken together, our results provide initial evidence that joint attention in ASL-based interactions is achieved when parents use a follow-in approach whereby they label objects to which the child had previously been attending. Deaf children learning ASL seek linguistic input related to their focus of attention through frequent and spontaneous gaze shifts to their interlocutors.

11:35

Break



Tuesday  
11:50

## Predictive processing of grammatical gender in bilingual children

Jasmijn Bosch & Francesca Foppolo

Listeners process speech incrementally and they predict what is coming ahead on the basis of lexical or morphosyntactic cues. One important morphosyntactic cue is grammatical gender, which three-year-old monolingual children use rapidly in spoken word recognition (Lew-Williams & Fernald, 2007). However, second language learners process grammatical gender less efficiently than native speakers (Lew-Williams & Fernald, 2010), and transfer effects have been reported when there is a mismatch in grammatical gender between two languages (Morales et al., 2016; Lemmerth & Hopp, 2019). We contribute to this debate by means of an online study on the processing of grammatical gender in German-Italian bilingual children, taking into account relative language proficiency.

We tested 40 German-Italian bilingual children between the ages of seven and nine (Mean Age = 8;5), using a visual world eye-tracking paradigm. Participants listened to Italian sentences starting with *Dov'è la/il ..?* ('Where is the ...?') in front of a visual scenario depicting two pictures while their eye-movements were recorded. Visual items either matched or mismatched in gender (*la fragola* 'the strawberry' vs. *la ciliegia* 'the cherry' or *il fungo* 'the mushroom') so that anticipatory eye-movements towards the noun could be observed at the article region in the predictable but not in the unpredictable condition. Moreover, the items varied with respect to their gender congruency in the two languages, so as to detect cross-linguistic influences in the anticipatory pattern (e.g. *mouse*, which is masculine in Italian and feminine in German vs. *butterfly*, which is feminine in Italian and masculine in German).

Hypothesizing that children integrate gender cues incrementally, we expected participants to anticipate looks at the target picture on the basis of the gender of the article. Hypothesizing that there is cross-linguistic influence, we further predicted this anticipation effect to be delayed when there is gender incongruency between German and Italian. Furthermore, we hypothesized that Italian-dominant children would show stronger anticipation than German-dominant children, while they were expected to be less likely to show cross-linguistic influence from German.

The data were analyzed by means of generalized linear mixed effect models on the proportion of fixations on the target as a function of condition, time region and relative language proficiency. First, a significant interaction between condition (predictable/unpredictable) and auditory region (intro/determiner) revealed an increase of looks to the target during the article in the predictable condition (Est. odds ratio=1.86, 95% CI=1.78..1.95,  $p < .0001$ ). Second, this effect was significantly stronger when gender was congruent in the two languages (Est. odds ratio=1.16, 95% CI=1.10..1.23,  $p < .0001$ ). Third, both effects seemed to be modulated by relative language proficiency; Italian-dominant children showed more anticipation ( $p < .001$ ) and less interference from German ( $p = .016$ ).

Our results provide evidence for rapid incremental processing of grammatical gender in bilingual children, although there may be cross-linguistic interference when the grammatical gender does not overlap across the two languages. Moreover, our findings suggest that language dominance matters; language processing may be more efficient when proficiency is greater, and the gender incongruency effect is most likely when participants are tested in their less dominant language.



Tuesday  
12:15

## How does bilingual reading guide object identification of visual referents?

Naomi Vingron, Madelaine Thomas & Debra Titone

When bilinguals read, they activate representations of word referents within each language (i.e., cross-language activation). This is evidenced by two well-documented effects during first (L1) and second language (L2) reading: 1) when words share visual form but not meaning (i.e., interlingual homographs, e.g., CRANE refers to a skull in French, but a machine in English); and 2) when words share both visual form and meaning across languages (i.e., cognates, e.g., LION refers to the animal in both languages). Of note, studies of bilingual reading differ from those of spoken word processing using the visual world paradigm, wherein people hear words and pick out actual objects (or pictures) in the physical world. Thus, in reading people typically imagine a word's visual referent; in visual world studies, people encounter actual visual referents in a display, a difference that likely modulates cross-language word activation over time depending on how easy it is to identify an object (Huettig, Olivers & Hartsuiker, 2011).

Consequently, it is unclear how written words guide object identification, and whether cross-language effects found in reading with imagined objects extend to reading with actual objects. To address this issue, we asked 49 bilingual adults (19 English/French; 30 French/English) to decide whether a visually presented word matched a visually presented image. On any given trial, participants first saw a word for 500 ms, and after a 150 ms delay, viewed a picture that either matched or mismatched the word. Participants completed an initial L1 block followed immediately by a L2 block. Each block included interlingual homographs, cognates, or language-unique words. The same participants then rated each image on visual complexity and label typicality.

There were three key findings. First, correct word-picture matching times increased along with object complexity for homographs, but not for frequency-matched language-unique words. This suggests that cross-language referential conflict was lessened for easily identifiable objects. Second, this finding did not occur for cognates, but was observed as part of an interaction with L2 exposure for language-unique words frequency-matched to the cognates (which were overall lower in frequency than language-unique words matched to homographs). Here, increased L2 exposure had no impact for low complexity objects, but led to longer correct word-picture matching times for high complexity objects. Finally, and replicating past work for sentence reading (e.g., Whitford & Titone, 2017), higher L2 exposure was associated with lower overall accuracy during L1 reading, but higher accuracy during L2 reading for many of the word-type conditions.

These results suggest that bilingual reading guides object identification of visual referents in a mutually constraining manner, particularly when word processing is difficult (i.e., for interlingual homographs, or language-unique words that are lower in frequency). Moreover, these processes are modulated by both visual complexity of the objects themselves, and individual differences in prior L2 experience. Taken together, these results are consistent with models of object-situated word processing (e.g., Huettig et al., 2011) suggesting that matching linguistic information to input from the environment is mediated by the availability of concept representations in memory.

12:40

Lunch served outside venue

Tuesday  
14:10

## The moment-to-moment pitch dynamics of child-directed speech shape toddlers' attention and learning

Mira Nencheva, Elise Piazza & Casey Lew-Williams

Toddlers prefer and learn from child-directed speech (CDS) over adult-directed speech (ADS) (Cooper & Aslin, 1990; Thiessen et al., 2005; Graf Estes & Hurley, 2013). However, little is known about how moment-to-moment features of CDS, such as variation in pitch over the course of a single word, drive real-time attention and learning. We extracted common word-level pitch contours from natural CDS and quantified how they affect toddlers' attention using pupil size synchrony.

In Study 1, using CHILDES corpora, we extracted pitch contours from natural CDS to one infant (6-12-m.o.) and two children (24-30-m.o.) (Soderstrom et al., 2008; Weist & Zevenbergen, 2008). Hierarchical clustering (Montero & Vilar, 2014) of noun pitch contours yielded 4 clusters: rises, falls, hills, and valleys.

In Study 2, we compared pupil size synchrony for CDS vs. ADS, and for the 4 CDS word-level contours. 24-30-month-olds ( $n=30$ ) listened to the same children's story once in CDS and once in ADS. For each trial, we extracted synchrony based on the pairwise dynamic time-warping distance (Tormene et al., 2008) between the pupil size time-series of the participants. When adults attend to particular moments in a stimulus, their pupil dilations synchronize, indicating shared attention (Kang & Wheatley, 2015, 2017). Consistent with increased attention to CDS, there was higher synchrony for CDS vs. ADS ( $p<0.001$ ). Intermixed with stories were 20 individual sentences with target nouns that followed the above-described contours, plus a flat baseline contour. We observed highest synchrony for hills, likely reflecting parents' pervasive use of this contour when referring to key words in CDS (Fernald & Mazzie, 1991; Aslin, 1993). Valleys and flats elicited the lowest synchrony, and rises and falls fell in-between. To validate these results in a naturalistic context, we clustered the pitch contours of words in the CDS story into contour types and quantified the synchrony during each word. Contour types shaped synchrony both in sentence trials ( $p<0.001$ ) and in natural stories ( $p<0.001$ ), with no significant interaction between contour type and word source ( $p\sim 0.7$ ), suggesting the same pattern in both contexts. Together these results show that CDS drives increased time-locked attention to speech in toddlers, and its specific subsecond pitch dynamics further shape attention in real time.

In Study 3, we examined toddlers' learning of novel words following contours that elicit higher (hills) vs. lower (valleys) synchrony. During training, toddlers heard sentences containing 4 novel words, each of which was consistently spoken in a hill or a valley contour and appeared with a unique isoluminant object. Toddlers learned novel words better when presented in high- vs. low-synchrony contours ( $p<0.05$ ), and moreover, synchrony during each individual novel word during training predicted learning of the same word during test ( $p<0.05$ ), suggesting that by modulating toddlers' attention, pitch dynamics in CDS impact learning.

This investigation yields a new, subsecond framework for understanding how toddlers engage with a signal known to support language learning. We identified the most common pitch contours in CDS and revealed a physiological response sensitive to real-time dynamics of attention and learning in response to these cues.

In natural listening conditions, there are massive fluctuations in listening demands for perceiving the speech signal accurately. Not only do listeners need to contend with different levels of background noise, but they also must resolve natural phonetic ambiguities in the speech stream. The degree of ambiguity (overlap) between phonetic categories varies significantly with the type of register a particular talker uses. For example, a talker speaking conversationally tends to articulate with less precision compared to when they talk in a clear speech register. If a talker produces the vowels in “bat” and “bet” nearly identically in conversational speech, this will create considerable competition between those respective phonetic categories that must be resolved by the listener to select the correct word. This overlap of phonetic categories in conversational speech leads to the hypothesis that listeners will expend greater effort to perceive conversational speech than clear speech. Attentional demands of perceiving conversational versus clear continuous speech are not well characterized, but pupillometry is a tool that can potentially probe online listening effort during continuous speech perception. This study used pupillometry to measure whether there is an attentional cost associated with passively listening to conversational speech compared to clear speech. If there are increased listening demands for speech with greater phonetic competition, we expect a steeper pupil dilation slope and a larger peak pupil dilation (PPD) compared to perceiving clear speech. Participants passively listened to sentences presented either in silence or to a challenging individually-calibrated signal-to-noise ratio (SNR) that corresponded to correctly perceiving 50% of the words in each sentence. Participants heard each sentence in both clear and conversational registers, with order counterbalanced across subjects. Each sentence was semantically incoherent (no extractable meaning) to prevent listeners from relying on context to resolve phonetic ambiguity. Intelligibility was measured with a visual word verification task after each sentence. As expected, behavioral results show a significant main effect of silence versus noise in accuracy and reaction time for responses to the visual word probe, but of interest, no interaction between clear and conversational speech in either behavioral measure. Nonetheless, pupil size curves indicate an interaction between noise and speech register. In silence, there is a higher and earlier PPD for conversational sentences than clear sentences, suggesting greater processing demands when phonetic ambiguity is high. The opposite is seen in noise—there was a greater PPD for clear sentences than for conversational sentences. These results indicate attention to the speech signal is allocated according to the amount of background noise and the degree of phonetic competition in the signal. Critically, there is a small attentional cost for perception of conversational speech even without competing background noise. This suggests that conversational speech imposes a processing penalty, even when behavioral data indicate equal intelligibility with clear speech.

We report a series of eye-tracking experiments in the visual-world paradigm in which we tested the hypothesis that the most prominent referent is more likely to be the perspectival center of a sentence or stretch of discourse compared to competing referents. As an indicator for perspective-taking we used free indirect discourse (FID), and as indicators of prominence we used grammatical function and type of referring expression. FID can only be interpreted correctly if the reader is able to take the protagonist's perspective.

In our first two experiments ( $n=40$  and  $n=35$ ) we tested the processing of FID compared to a neutral story continuation in order to establish the visual world paradigm as a measure for perspective shifts. We found an increase of gazes on the subject of a sentence in the FID condition compared to the neutral story condition as a result of the change towards the subject's perspective triggered by the FID starting at 1500ms from the onset of the target sentence. We regard this increase of gazes on the subject to be the result of the ascription of authorship of the utterance in FID.

In a follow up experiment ( $n=42$ ) we investigate the role of prominence lending cues on perspective-taking. Our test items consist of short stories introducing two referents, one in subject position and one in object position in the first sentence. In order to test for the role of prominence of the protagonists we manipulated grammatical function (subject or object) and referring expression (proper name or indefinite NP) in the first sentence. We predicted that - due to prominence hierarchies - subjects are more available as anchors for FID than objects and proper names are more available as anchors for FID than indefinites. We plotted the shares on the subject and the object during the first 3500ms from the onset of the target sentence expecting to find an effect starting at about 1500ms. When referred to with the same type of expression the proportion of gazes on the subject increases faster than gazes on the object suggesting that grammatical function has an impact on the availability as the anchor for FID. Regarding the different referring expressions in subject position with a fixed referring expression for the object we observe a weak difference. However, referring expression results in a higher share of gazes when there is a maximal contrast in prominence, i.e. when the subject is referred to with a proper name while the object is referred to with an indefinite NP and vice versa. We modeled the data using multinomial mixed models with the conditions and 15 equally spaced time points as predictors. Due to a high variation of individual items and participants the differences remained statistically uncertain. Yet, we regard these results as promising insights for further research employing the visual-world paradigm for investigating perspective-taking in language processing.

83 Chinese participants first finished a screen-based puzzle game with a computer player. Half of the participants were led to believe that they were playing with a real human participant. In the high-coordination condition, the computer player presented participants with puzzle pieces that could be placed near their partner's last placed piece. In the low-coordination condition, participants were presented pieces that could only be placed further away from their partner's last placed piece.

We then measured participant's performance in a perspective taking task adapted from Heller et al. (2008). Participants heard auditory instructions from the previous computer partner such as "Please give me the big cubic block" in Chinese and responded to the request by clicking on the corresponding block. The display included two pairs of size-contrasting blocks, e.g. a big cubic block (target) and a small cubic block (target-contrast), and a big triangular cone (competitor) and a small triangular cone (competitor-contrast). Two conditions were constructed for the factor Ground: the shared condition (the competitor-contrast object is visible to both the participant and the partner) and the privileged condition (the competitor-contrast object is only visible to the participant). Participants' eye-movements were recorded. If participants were taking into account the speaker's perspective, then there should be more looks to the target when the target-competitor is privileged.

We performed a growth curve analysis (Barr, 2008) on the proportion of looks to the target during the critical time region (the scalar adjective region). There is a main effect of Ground: the proportion of looks at the target in the privileged condition is higher than those in the shared condition ( $\beta=0.312$ ,  $SE=0.037$ ,  $p<0.001$ ), replicating the main findings of the Heller et al. study, providing further evidence that perspective influences real-time comprehension. There are also interaction effects of Ground\*Coordination\*Time ( $\beta=-0.717$ ,  $SE=0.249$ ,  $p=0.004$ ) and Ground\*Coordination\*Time<sup>3</sup> ( $\beta=2.295$ ,  $SE=0.778$ ,  $p=0.003$ ), indicating that the growth curve patterns of the proportion of looks differ across different Ground\*Coordination groups (Fig.1). For the high-coordination group, the proportion of looks to the target in the privileged ground condition peaked earlier than those of the low-coordination group. This suggests that, with high-coordination experience, participants were more attuned to their partner's perspective.

These results show that fine-grained coordination improves perspective taking in online communication, suggesting that joint actions play an important role in facilitating social cognitive tasks.

Tuesday  
16:10

**Empathy influences behavioral perceptions and eye movements in non-literal language processing**

Pavitra Makarla, Lauren Benson, Hana Kim, Gitte Joergensen & Kathrin Rothermich

Recognizing nonliteral language is a crucial part of communicative functioning and involves complex social cognitive skills such as the inference of mental states and empathy. To study the influence of empathy on processing literal positive, blunt, sarcastic, and jocular dynamic social interactions, we tracked healthy adults' eye movements while they watched video vignettes (N = 40). Participants were asked to evaluate speaker intention and friendliness, and we measured their empathy levels using the Interpersonal Reactivity Index (Davis, 1980). Behavioral results showed that participants with higher levels of empathic concern rated blunt interactions as less friendly compared to literal positive interactions, while participants with low empathic concern levels showed the opposite pattern. The eye tracking analysis revealed that all participants spent significantly more time looking at faces when scanning literal versus nonliteral vignettes. Moreover, participants with higher empathic concern showed significantly more fixations to faces when viewing jocular interactions compared to participants with lower empathic concern scores. Our findings will be discussed not only in the context of interactive versus modular nonliteral language processing theories but also in relation to clinical populations who exhibit difficulties with empathy and nonliteral language understanding, (e.g. Autism Spectrum Disorders or Parkinson's Disease).

16:35

**Break**

Tuesday  
16:50

**KEYNOTE**

The nascent lexicon: Comprehension in the lab and in the world

Elika Bergelson

*Abstract forthcoming*

17:55

Break

18:15

Dinner served in Bousfield



<b>WEDNESDAY, MARCH 18</b>	
8:15	<b>Breakfast</b>
<b>KEYNOTE</b>	
9:00	<b>Intention, attention, and real-time language processing in little listeners</b> Craig Chambers
<i>Abstract forthcoming</i>	
9:55	<b>The timecourse of spoken word recognition across development</b> Elizabeth Schoen Simmons & James S. Magnuson
<p>The fine-grained timecourse of adult spoken word recognition (SWR) has been studied extensively, but there have been few studies with young children. A key finding with adults is that words that are similar at onset (cohorts) compete early and strongly, while rhymes compete later and more weakly. This is intriguing developmentally, because multiple theories predict that young (pre-reading) children may not show rhyme competition. We created a simplified visual world paradigm for typically developing pre-reading children and early readers. Participants heard an auditory instruction and used a mouse to click on the named picture. Targets were paired with a cohort competitor, rhyme competitor or phonologically unrelated object. Even our youngest participants demonstrated phonological competition similar to that of adults. Our findings suggest that pre-readers have access to differentiated lexical representations and utilize subtle, temporal differences in the signal to process spoken words.</p>	
10:20	<b>Break</b>

Wednesday  
10:40

## The role of early language exposure and modality on bottom-up visual attention.

Allison Fitch, Sudha Arunachalam & Amy Lieberman

Deaf individuals demonstrate superior visual abilities relative to hearing individuals on a range of tasks. This may be attributed to experience with sign language or to neural plasticity associated with deafness. Findings from spatial orienting and disengagement, in which deaf participants are faster to make eye movements toward peripheral stimuli than hearing participants, suggest it is the latter (Bosworth & Dobkins, 2002; Colmenero et al., 2004).

However, these findings assess only top-down attentional orienting, in which stimuli are voluntarily and selectively attended to (Chica et al., 2013). Bottom-up attentional orienting, in contrast, is an involuntary process in which stimuli automatically capture attention. Sign language exposure provides an environment rich in bottom-up cues including visual attention-getting (e.g. waving) and signing in the visual field. Early exposure to this linguistic environment may “train” the underlying system that supports spatial orienting during development. To test this, we conducted a bottom-up orienting task (Gap-Overlap Paradigm; adapted from Fisher et al., 2016) to determine the role of early sign experience and deafness on spatial orienting and disengagement.

Forty-eight adults from three groups participated: Deaf signers exposed to ASL prior to age 5 ( $n = 18$ ), Deaf signers exposed to ASL after age 5 ( $n = 18$ ), and hearing non-signers ( $n = 12$ ). Orienting and disengagement were measured by comparing saccadic reaction times (SRT) to peripheral stimuli in trials that did and did not require disengagement from a central stimulus. Each trial began with a central fixation cross, which was replaced by a central stimulus. In Shift trials, the central stimulus disappeared prior to the onset of the peripheral stimulus, 500 to 1500 ms later. Disengage trials were the same, except the central stimulus remained at central fixation throughout the duration of the trial. Participants thus needed to disengage their attention from the central stimulus in order to orient to the peripheral stimulus.

As expected, participants were slower to orient to the peripheral stimulus on Disengage, relative to Shift trials,  $F(1,46) = 54.2$ ,  $p < .001$ . Group-level differences were not statistically significant,  $F(2,46) = 1.2$ ,  $p = .3$ . Additionally, there were no interactions between group and trial type. However, for the Deaf participants, age of ASL acquisition was related to Shift SRT ( $r = .34$ ,  $p = .048$ ), but not Disengage SRT ( $r = .26$ ,  $p = .13$ ).

Preliminary findings suggest that for Deaf individuals, early acquisition of a sign language promotes orienting in a bottom-up attention task. This differs from the literature on top-down attention, in which deaf individuals are faster than hearing individuals regardless of language experience. The exact role early language acquisition plays in gaze shifting remains to be tested. We suggest that early language acquisition may be a proxy for experience in high-quality joint attention interactions, which tend to be reduced for deaf individuals who are later acquirers of ASL (Frischen et al., 2007). These interactions are likely to be rich with bottom-up cues that may be sufficient to train the underlying visual attention mechanisms that promote orienting.

Wednesday 11:05	<b>Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing</b>  Falk Huettig & Ernesto Guerra
<p>There is a consensus among language researchers that people can predict upcoming language. But do people always predict when comprehending language? Notions that “brains ... are essentially prediction machines” certainly suggest so. In three eye-tracking experiments we tested this view. Participants listened to simple Dutch sentences (“Look at the displayed bicycle”) while viewing four objects (a target, e.g. a bicycle, and three unrelated distractors). We used the identical visual stimuli and the same spoken sentences but varied speech rates, preview time, and participant instructions. Target nouns were preceded by definite gender-marked determiners, which allowed participants to predict the target object because only the targets but not the distractors agreed in gender with the determiner. In Experiment 1, participants had four seconds preview and sentences were presented either in a slow or a normal speech rate. Participants predicted the targets as soon as they heard the determiner in both conditions. Experiment 2 was identical except that participants were given only a one second preview. Participants predicted the targets only in the slow speech condition. Experiment 3 was identical to Experiment 2 except that participants were explicitly told to predict. This led only to a small prediction effect in the normal speech condition. Thus, a normal speech rate only afforded prediction if participants had an extensive preview. Even the explicit instruction to predict the target resulted in only a small anticipation effect with a normal speech rate and a short preview. These findings are problematic for theoretical proposals that assume that prediction pervades cognition.</p>	
11:30	<b>Lunch</b>
<b>KEYNOTE</b>	
12:30	<b>Perception, Action and Language with Virtual Reality: Low hanging fruit or deluded speculations?</b>  Michael Tanenhaus
<p>Some of the rationale for the task-based visual world research initiated by my lab more than 25 years ago was to leverage the combination of a visual workspace, a well-defined task, and spoken language to ask questions about linguistic, perceptual and conceptual representations that would be otherwise difficult to ask. While 2-D screen-based paradigms have advantages in stimulus control and are thus well-suited to addressing some classes of questions, they are not well-suited to others. For example, Craig Chambers work on action-contingent affordances in circumscribing referential domains requires real actions and three-dimensional objects. In this talk, I will outline a vision for a research program that would use virtual reality, along with saccade and action-based screen changes, to address questions about how aspects of representations might be distributed among perceptual, motor, and linguistic representations, rather than integrated into a central conceptual representation. I'll moot some hypotheses, and propose some technologically demanding, but increasingly feasible, lines of research. This is work I won't do: I've transitioned to retirement, no longer have an active research program, and am increasingly less conversant with the relevant literature. However, I hope to make enough of a case to encourage some of you to think about, and perhaps even pursue, some of these issues.</p>	
13:20	Closing remarks
13:40	Bus leaves Nathan Hale for UMass-Amherst