How common ground information affects reference resolution in children and adults: evidence from behavioral data, eye-tracking and ERPs

CoGCI, University of Potsdam
Maria Richter, Choonkyu Lee, Mariella Paul, Isabell Wartenburger and Barbara Höhle

In communication, ambiguous referring expressions which have multiple referents within the linguistic or non-linguistic context can be resolved by taking another person's perspective. This requires considering so-called common ground (CG) information, that is, mentally or perceptually shared information of both interlocutors. Competing accounts have been proposed about when CG information is integrated during utterance comprehension. In the *Constraint-Based Account*, listeners can make rapid use of CG during their online utterance interpretation (e.g., Nadig and Sedivy, 2002; Hanna et al., 2003). In contrast, the *Egocentrism Account* assumes that CG information is integrated only late with effort, after an initial egocentric interpretation has led to errors (e.g., Keysar et al., 2000). The *Autonomous Activation Account* instead suggests that listeners actively attempt to take a speaker's perspective in anticipation of a linguistic expression, but then fail to fully integrate CG information due to autonomous activation of privileged information (Epley et al., 2004; Barr, 2008). Since it is still a matter of debate, when CG information is integrated during utterance comprehension, we further investigated this topic from two perspectives. On the one hand, we were interested in the time course of WHEN CG information affects utterance processing in adults by means of eye-tracking as well as electroencephalography (EEG). The use of EEG resembles a novel approach in order to examine the temporal dynamics of CG integration. Further, it allows for the investigation of the electrophysiological underpinnings of CG processing through analyzing event-related potentials (ERPs) and time-frequency data (TFA). On the other hand, we considered the question of WHEN from a developmental perspective: we examined 4- and 5-year-old children with the visual-world eye tracking paradigm to specify the developmental trajectory of adult-like sensitivity to CG information in online processing.

Both children and adults played largely identical versions of the referential communication game (e.g., Keysar et al., 2000). The crucial feature of this game is to manipulate the conditions of visual access between interlocutors by occluding certain objects in a grid: some objects are in CG (i.e., visible for both interlocutors), to some objects only the participant has privileged access to.

For both the eye-tracking and the EEG study in adults, participants (eye-tracking: n=27, EEG: n=34) saw a 4x4 grid containing two sets of three different sized objects (i.e., small, medium, big) and two single distractors on a screen. A virtual confederate behind the grid provided auditory instructions (e.g., "Move the big star to the top."). Participants then had to drag and drop the corresponding object to a field above the grid using a computer mouse. In conflict trials, the object that fit the request best from the perspective of the participant ("competitor"; e.g., "the big star") was occluded from the confederate's view. Thus, participants had to consider CG information to select the correct object ("target"; e.g., the medium-sized star). The experiment also entailed a no-conflict condition (targets visible for confederate and participant), a no-hidden condition (no occlusions at all), and filler trials. In advance of the experimental phase participants were instructed to mind the confederate's perspective and underwent a practice phase with corrective feedback (2-step instruction similar to Wang et al., 2016). During the test phase, eye-tracking or EEG data were recorded, and reaction times (RTs) and accuracy rates were measured.

The accuracy results revealed that participants mastered integration of CG in the conflict condition (eye-tracking: 99,6%, EEG: 100% accuracy rate). However, perspective taking had its costs: RTs showed that participants were on average slower in the conflict condition compared to the no-conflict condition (eye-tracking: 219ms, EEG: 214ms). The increased processing costs for the conflict vs. no-conflict condition were also reflected in the eye-tracking as well as in the ERP and TFA data as revealed by cluster-based permutation analyses (Maris and Oostenveld, 2007).For the eye-tracking data, later (cluster 400-850ms relative to noun onset) and longer (cluster 1500-1800ms relative to noun onset) looks to target were revealed in the conflict condition in comparison to the no-conflict condition. For the ERP data, the conflict

condition elicited an enhanced late positivity in posterior (800-850ms relative to noun onset) and anterior (850-950ms and 1050-1150ms relative to noun onset) brain regions in comparison to the no-conflict condition. TFA data are sensitive to power changes in the frequency-bands and showed a power increase for the conflict vs. no-conflict conditions in theta (4-7 Hz) and alpha bands (8-12 Hz) 400-1050ms relative to noun onset.
In sum, all of the applied methods revealed processing differences between the conflict and no-conflict conditions. The differences starting around 400ms relative to the onset of the noun were signified by later and longer looks to target (eye-tracking), enhanced late positivities (ERPs) and a power increase mostly visible in the theta frequency band (TFA). Since late positivities index processes aligned to the modification of discourse representations as well as conflict resolution (e.g., van Herten et al., 2005; Bornkessel-Schlesewsky and Schlesewsky, 2008; Schumacher, 2009; McCleery et al., 2011), and since increased power in the theta band has been found to be associated to slow potentials in the ERPs (Herrmann et al., 2005), reflecting, for instance, the processing of complex events (Başar et al., 2001),the integration of CG information seemed to be an effortful and relatively late process (i.e., 400ms after the presentation of the critical noun). Our findings therefore rather support the Egocentrism Account (Keysar et al., 2000), or the Autonomous Activation Account (Barr, 2008), rather than the Constraint-Based Account (Nadig and Sedivy, 2002; Hanna et al., 2003).This is especially remarkable in our setting, as clear instructions were given, the repetition rate was high, and the responses were perfectly accurate.

In the developmental literature, there is a similar conflict between studies that find immediate online effects of CG information on eye-movements in children (Nadig and Sedivy, 2002) and those that demonstrate autonomous activation of privileged information in eye-movements in both children and adults (Epley et al., 2004). In a parallel study with four- and five-year-old children, we investigated the developmental trajectory of adult-like sensitivity to CG information in online processing of reference. In this study, we presented pre-recorded videos of a 3x3 grid and an adult confederate making requests, such as “Pick up the small horse”, to the participant. We also tracked children’s eye-movements to find out whether and at what point in the sentence processing children take CG information into account. In a preliminary analysis, the gaze data showed that, in the conflict condition, looks to the hidden competitor increased dramatically around the beginning of the critical size or position adjective for most participants (from 3% at 600ms after the auditory onset to 34% at 1700ms), while looks to the target in common ground increased much less overall (from 8% to 14% in the same interval). In the no-conflict condition, in contrast, looks to the target in common ground increased dramatically (from 11% at 600ms to 42% at 1700ms), while those to a hidden object hardly increased at all (from 11% to 13%). Analysis of the behavioral responses from our pilot study with real-life interaction with an adult speaker showed that both age groups mostly ignored the speaker’s perspective (17 out of 30 failed all critical trials), although 5-year-olds had a statistically non-significant tendency to choose the non-privileged object more often (40%) than 4-year-olds (20%). In sum, children up to about 6 years of age have difficulty taking an interlocutor’s visual perspective while determining the relevant referential domain in general, even though older children showed a tendency toward developing this ability. These results support Epley et al.’s (2004) account in which autonomous activation of privileged information is difficult to suppress especially for children even after training on the use of CG information.

References

Barr, D.J., 2008. Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground. Cognition 109 (1), 18–40. 10.1016/j.cognition.2008.07.005.

Başar, E., Başar-Eroglu, C., Karakaş, S., Schürmann, M., 2001. Gamma, alpha, delta, and theta oscillations govern cognitive processes. International Journal of Psychophysiology 39 (2-3), 241–248. 10.1016/S0167-8760(00)00145-8.

Bornkessel-Schlesewsky, I., Schlesewsky, M., 2008. An alternative perspective on "semantic P600" effects in language comprehension. Brain research reviews 59 (1), 55–73. 10.1016/j.brainresrev.2008.05.003.

Epley, N., Morewedge, C.K., Keysar, B., 2004. Perspective taking in children and adults: Equivalent egocentrism but differential correction. Journal of Experimental Social Psychology 40 (6), 760–768. 10.1016/j.jesp.2004.02.002.

Hanna, J.E., Tanenhaus, M.K., Trueswell, J.C., 2003. The effects of common ground and perspective on domains of referential interpretation. Journal of Memory and Language 49 (1), 43–61. 10.1016/S0749-596X(03)00022-6.

Herrmann, C.S., Grigutsch, M., Busch, N.A., 2005. EEG oscillations and wavelet analysis, in: Handy, T.C. (Ed.), Event-Related Potentials. A Methods Handbook. MIT Press, Cambridge, Massachusetts.

Keysar, B., Barr, D.J., Balin, J.A., Brauner, J.S., 2000. Taking Perspective in Conversation: The Role of Mutual Knowledge in Comprehension. Psychological Science 11 (1), 32–38. 10.1111/1467-9280.00211.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. Journal of neuroscience methods 164 (1), 177–190. 10.1016/j.jneumeth.2007.03.024.

McCleery, J.P., Surtees, Andrew D R, Graham, K.A., Richards, J.E., Apperly, I.A., 2011. The neural and cognitive time course of theory of mind. The Journal of neuroscience : the official journal of the Society for Neuroscience 31 (36), 12849–12854. 10.1523/JNEUROSCI.1392-11.2011.

Nadig, A.S., Sedivy, J.C., 2002. Evidence of Perspective-Taking Constraints in Children's On-Line Reference Resolution. Psychological Science 13 (4), 329–336. 10.1111/j.0956-7976.2002.00460.x.

Schumacher, P.B., 2009. Definiteness Marking Shows Late Effects during Discourse Processing: Evidence from ERPs, in: Devi, S.L., Branco, A., Mitkov, R. (Eds.), DAARC 2009. LNAI 5847. Springer-Verlag, Berlin Heidelberg, pp. 91–106.

van Herten, M., Kolk, H.H.J., Chwilla, D.J., 2005. An ERP study of P600 effects elicited by semantic anomalies. Brain research. Cognitive brain research 22 (2), 241–255. 10.1016/j.cogbrainres.2004.09.002.

Wang, J.J., Ali, M., Frisson, S., & Apperly, I.A. (2016). Language complexity modulates 8- and 10-year-olds' success at using their theory of mind abilities in a communication task. Journal of Experimental Child Psychology, 149, 62-71.