

SSLP 2018 program

Tuesday, 4th September

Time	Topic
8.30 – 9.15	Welcome coffee & badge pick-up
9.15 – 9.30	Opening remarks
9.30 – 10.30	Keynote: Jos van Berkum
10.30 – 10.45	Short coffee break
10.45 – 11.15	<i>The impact of stereotypes and noun endings on processing gender in English: comparing native and non-native performance</i> (J. Müller, L. Konieczny, V. Haser)
11.15 – 11.45	<i>Empathy determines how intonation is used to process semantically ambiguous word</i> (N. Esteve-Gibert, A. Schafer, B. Hemforth, C. Portes, C. Pozniak, M. D'Imperio)
11.45 – 12.15	<i>Emotion, reference, and perspective-taking in children's language processing</i> (C. G. Chambers, J. Berman, M. Khu, J. Thacker, S. A. Graham)
12.15 – 13.45	Lunch break
13.45 – 14.45	Keynote: Stefanie Jannedy
14.45 – 16.00	Poster* session with coffee
16.00 – 16.30	<i>Do posh ducks say qu[ɑ:]ck?: Investigating the cognitive representation of dialect variation</i> (M. Austen)
16.30 – 17.00	<i>Automatic vs. RT-modulated phonetic convergence over timescales and context</i> (S. J. Tobin, M. A. Hullebus, A. I. Gafos)
17.00 – 17.30	<i>Processing of case variation in German prepositional phrases</i> (A. Engel & A. Hanulíková)
19.00 – 23.00	Social event: boat trip with small picnic

*Posters:

1. *Reference resolution and the integration of referential visual cues* (M. Sekicki & M. Staudte)
2. *The difference is not orthographic, it is grammatical: Situational variation of linguistic explanations* (M. Kuehnast & B. Lütke)
3. *Production/Interpretation Asymmetries in French Grammatical Gender* (Célia Richy & Heather Burnett)
4. *Toward an understanding of how nonbinary pronouns are learned and processed* (L. Ackerman)
5. *Tailoring Referential Descriptions: Talking to Children, Adults, and Computers* (R. Saryazdi, J. Bannon & C. G. Chambers)
6. *Integration and anticipation processes of the speaker and meaning in adults with and without autism spectrum disorder: evidence from eye-tracking and ERPS* (M. Barzy, J. Black, D. Williams, H. J. Ferguson)

Wednesday, 5th September

Time	Topic
8.30 – 9.00	Coffee
9.00 – 10.00	Keynote: Laura Staum Casasanto
10.00 – 10.30	Coffee break
10.30 – 11.00	<i>Integrating listener and speaker characteristics into the Coordinated Interplay Account</i> (K. Münster & P. Knoeferle)
11.00 – 11.30	<i>Integrating socially situated non-linguistic cues in pragmatic generalization</i> (A. Pogue, S. Brown-Schmidt, C. Kurumada)
11.30 – 12.00	<i>Effects of talker identity on speech comprehension across the lifespan</i> (A. Hanulíková)
12.00 – 13.30	Lunch break
13.30 – 14.30	Keynote: Shiri Lev-Ari
14.30 – 15.00	Coffee break
15.00 – 15.30	<i>Novel lexical representations are shaped by speakers' in-group status and learners' in-group biases</i> (S. Iacozza, A.S. Meyer, S. Lev-Ari)
15.30 – 16.00	<i>Do these pants make me look fat? Influences of response delay on conversational meaning</i> (A. Baltaretu & C. G. Chambers)
16.00 – 16.30	<i>Variation in French partial interrogatives: social meaning as a key factor to understand sociolinguistic norm violations</i> (G. Thiberge & B. Hemforth)
16.30 – 17.00	Panel discussion & concluding remarks

Talk abstracts – 4th September

THE IMPACT OF STEREOTYPES AND NOUN ENDINGS ON PROCESSING GENDER IN ENGLISH: COMPARING NATIVE AND NON-NATIVE PERFORMANCE

Julia Müller, Lars Konieczny & Verena Haser (all affiliated with
Albert-Ludwigs-Universität Freiburg)
julia.mueller@pluto.uni-freiburg.de

This study aims at establishing whether the perception of the noun ending *-er* as masculine carries over into English for native speakers of German. Since German is a language with a grammatical gender system (Hellinger & Bußmann, 2003), professions which end in *-er* identify a person as male in German (e.g. *der Gärtner* – “the (male) gardener”) while in English, they are not grammatically marked for gender. Thus, it is possible that Germans might still perceive professions ending in *-er* in English as more male. Alternatively, they might switch to an English native speaker view and judge professions based on their stereotypical associations instead, a phenomenon that has been established in various studies (e.g. Gygax, Gabriel, Sarrasin, Oakhill, & Garnham, 2009).

Participants read short texts while their eye movements were recorded. 60 stimuli with three sentences each were constructed. The first sentence introduced an occupational noun which varies in its stereotypical association (with males or females) between conditions. The third and final sentence then either included “men” and “women”, referring back to the occupational noun in the first sentence. For example: “During the last month, the stockbrokers/hairdressers tried to get the business going. Recently, it had gotten a bad reputation. But two of these men/women had a brilliant idea that would turn the ship around.”

Professions were taken from Misersky et al. (2014) who had native speakers of seven languages rate them for stereotypicality. Beyond that, professions ending in *-er* in English were contrasted with those that ended in *-or* and those that had endings which rarely occur in German (such as *-ian*). Additionally, 24 items from another reading experiment were used as distractors.

Beyond the stereotypical associations, gender, noun endings, and native languages, the participants’ proficiency and acquisition environment of English was established. Furthermore, they filled in a short form of Bem’s Sex Role Inventory (Bem, 1981), two surveys that measure sexist attitudes towards women (the Ambivalent Sexism Inventory, Glick & Fiske, 1996), and men (the Ambivalence Toward Men Inventory, Glick & Fiske, 1999), and the Modern Sexism Scale (Swim, Aikin, Hall, and Hunter, 1995).

Results from 64 participants (40 German, 24 English) indicate that *-er* may slightly, but *-or* more considerably, slow down processing for German speakers when used with *women*. Moreover, the increased exposure to incongruent stereotype-person combinations seemed to facilitate the acceptance of people in professions which do not correspond to the stereotype.

References

- Bem, S. L. (1981). *Bem Sex-Role Inventory: Professional manual*. Palo Alto, CA: Consulting Psychologists Press. Retrieved from <http://worldcatlibraries.org/wcpa/oclc/9335096>
- Glick, P., & Fiske, S. T. F. (1996). The Ambivalent Sexism Inventory: Differentiating Hostile and Benevolent Sexism. *Journal of Personality and Social Psychology*, *70*, 491–512.
- Glick, P., & Fiske, S. T. (1999). The Ambivalence Toward Men Inventory. *Psychology of Women Quarterly*, *23*(3), 519–536. <https://doi.org/10.1111/j.1471-6402.1999.tb00379.x>
- Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2009). Some grammatical rules are more difficult than others: The case of the generic interpretation of the masculine. *European Journal of Psychology of Education*, *14*(2), 235–246.
- Hellinger, M., & Bußmann, H. (Eds.). (2003). *Gender across languages: Volume III. Impact: Studies in language and society*. Amsterdam: Benjamins.
- Misersky, J., Gygax, P. M., Canal, P., Gabriel, U., Garnham, A., Braun, F., . . . Sczesny, S. (2014). Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior Research Methods*, *46*(3), 841–871. <https://doi.org/10.3758/s13428-013-0409-z>
- Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology*, *68*(2), 199–214. <https://doi.org/10.1037/0022-3514.68.2.199>

Empathy determines how intonation is used to process semantically ambiguous words

Núria Esteve-Gibert^{1,2}, Amy Schafer³, Barbara Hemforth⁴, Cristel Portes¹, Céline Pozniak⁴,
Mariapaola D'Imperio¹

¹Aix Marseille Univ, CNRS, LPL; ²Universitat de Barcelona; ³University of Hawai'i at Mānoa;

⁴Université Paris Diderot, CNRS, LLF

This study investigates how social skills impact on the processing of prosodic cues when the meaning of words is temporally ambiguous. The meaning of an utterance comes from the lexical-semantic value of words but also from its prosodic features (like intonation) or the sentence and situational context in which the utterance is framed. Individual pragmatic skills have been found to determine how listeners process these various sources of linguistic meaning: pragmatically-skilled individuals prefer non-literal interpretations [1-3] and consider all possible meaning alternatives [4], and individuals with higher perspective-taking abilities are more sensitive to prosodic prominence [5].

Using a visual-world eye-tracking task, we examined how individual empathic skills determine the processing of prosodic and contextual cues to disambiguate semantically ambiguous words. French-speaking participants (N=54) saw a virtual card game in which speaker A had to guess the image on speaker's B card, and then speaker B confirmed or contradicted speaker's A guess. Tests trials were 18 sentence pairs containing a suggestion (e.g. *Je pense que tu as un/e CW* 'I think you have a/an CW') and a response (*J'ai un/e CW* 'I have a/an CW'), where CW was the critical word and a homophone. The visual scene of the suggestion depicted the less-frequent alternative of the homophone. The visual scene of the response displayed 4 images: the suggested less-frequent alternative of the homophone, its more-frequent alternative, and 2 images of non-homophonic words. Responses were of 3 types: (a) confirmation response, with a falling intonation, (b) contrast response, with a rise-fall intonation, and (c) control response, with rise-fall intonation and a non-homophonic CW (Table 1). Additionally, 36 filler pairs contained non-homophonic CW in both the suggestion and the response. Participants also filled in the French version of the Empathy Quotient questionnaire [6].

We calculated fixation proportions to the image depicting the suggested word (the target in the confirmation responses; the competitor in the contrast and control responses) out of looks to all areas of interest, as a function of the listeners' empathic skills (2 groups: high empathy or HE, low empathy or LE). Figure 1 shows that during the processing of the CW, LE individuals look at the image depicting the suggested word equally in the 'confirmation response' and the 'contrast response'. However, the looking pattern of HE individuals during the processing of the CW differed between the confirmation-response and the contrast-response: they initially look more at the suggested image in the confirmation-response trials. As expected, all participants stopped looking at the suggested image during the processing of the non-homophonous CW in the control-response. Mixed-effects linear regression models on empirical logits confirmed these patterns, revealing a main effect of empathy ($\beta = 0.121$, $t = 2.386$, $p < .05$) and of prosody ($\beta = 0.127$, $t = 2.465$, $p < .05$), and a marginal interaction between the two ($\beta = -0.125$, $t = -1.859$, $p = .06$) during the 450-675ms time interval, corresponding to the processing of the second half of the CW. A post-hoc offline task was designed to understand the empathy*prosody interaction during the CW processing. It revealed that HE individuals did not form a different intonation-meaning mapping, but that they used prosodic information earlier in the utterance (pitch accent) to build up an intonation-meaning hypothesis. This study points at the need for taking into account individual social and pragmatic skills when investigating the online processing of intonation and meaning.

Condition	Speaker's A suggestion	Speaker's B response
Confirmation response	Je pense que tu as un/e [CW _{HOM}] <i>Je pense que tu as une <u>cane</u></i> 'I think you have a female duck'	J'ai un/e [CW _{HOM-SAME}] <i>J'ai une <u>cane</u> ^{LHIL* L-L%}, bien sûr, ...</i> 'I have a female duck, indeed, ...'
Contrast response	Je pense que tu as un/e [CW _{HOM}] <i>Je pense que tu as une <u>cane</u></i> 'I think you have a female duck'	J'ai un/e [CW _{HOM-DIFFERENT}] <i>J'ai une <u>canne</u> ^{LH* L-L%}, plutôt, ...</i> 'I have a stick, instead, ...'
Control response	Je pense que tu as un/e [CW _{HOM}] <i>Je pense que tu as une <u>cane</u></i> 'I think you have a female duck'	J'ai un/e [CW _{NONHOM}] <i>J'ai une <u>poupée</u> ^{LH* L-L%}, plutôt, ...</i> 'I have a doll, instead, ...'

Table 1. Summary and examples of all test conditions.

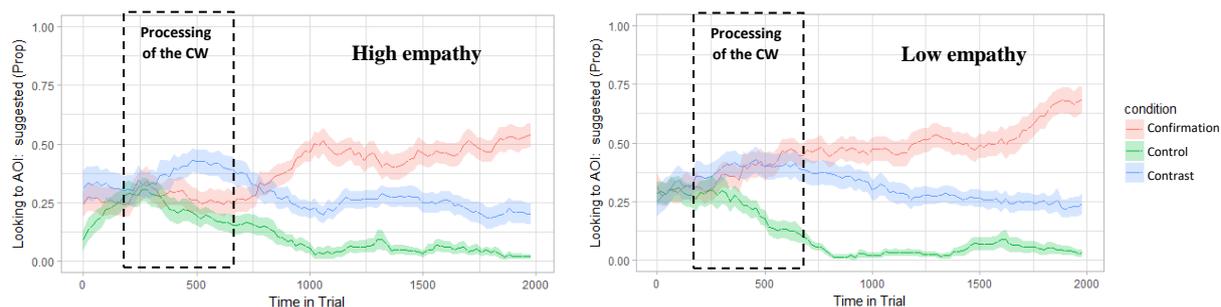


Figure 1. Proportion of fixations to the suggested image across conditions and time (ms), averaged across items, aligned to the onset of the critical word (CW). Because eye movements launched within 200ms of a stimulus onset are assumed to be influenced by that stimulus, the dashed rectangle signals the time interval in which listeners processed the critical word (CW). Shaded regions reflect 95% confidence intervals.

References

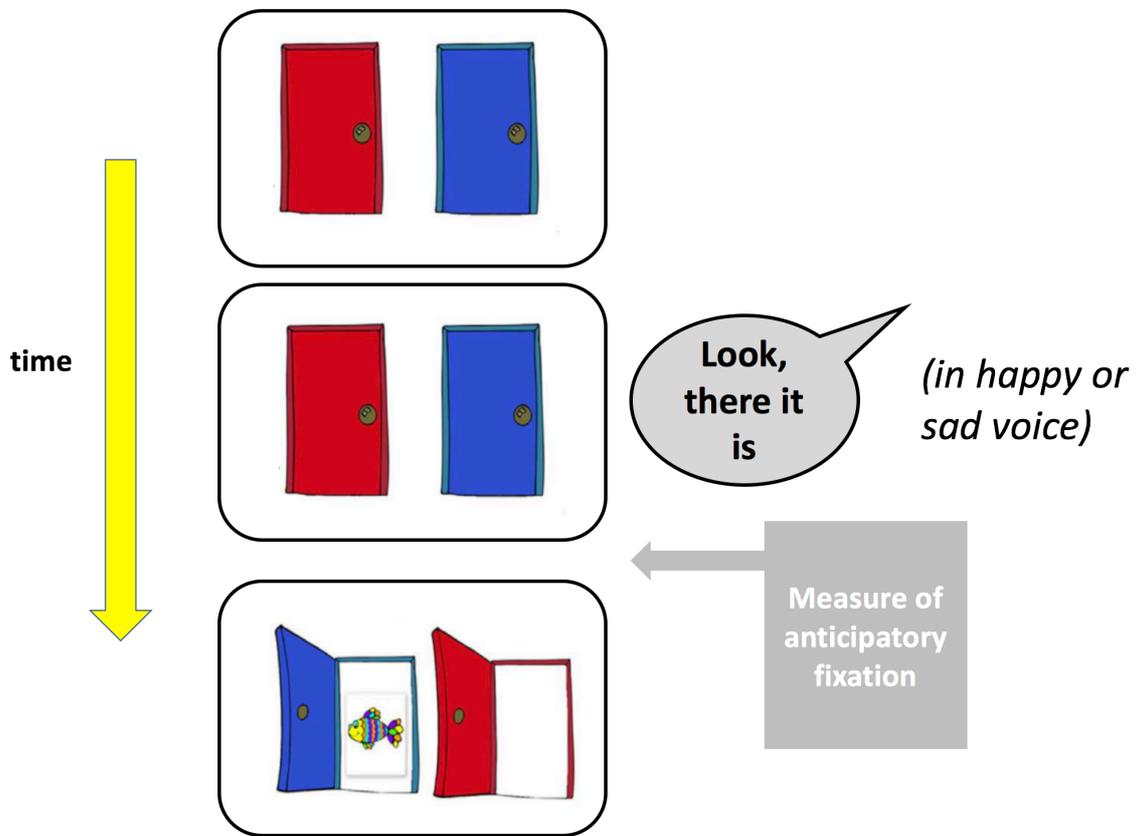
- [1] Li, S., Jiang, X., Yu, H., & Zhou, X. (2014). Cognitive empathy modulates the processing of pragmatic constraints during sentence comprehension. *Social Cognitive and Affective Neuroscience* 9, 1166–1174.
- [2] Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language*, 63(3), 324–346.
- [3] Zhao, M., Liu, T., Chen, G., & Chen, F. (2015). Are scalar implicatures automatically processed and different for each individual? A mismatch negativity (MMN) study. *Brain Research*, 1599, 137–149.
- [4] Degen, J., & Tanenhaus, M. K. (2015). Availability of alternatives and the processing of scalar implicatures: A visual world eye-tracking study. *Cognitive Science*, 1–30.
- [5] Jun, S., & Bishop, J. (2014). Implicit Prosodic Priming and Autistic Traits in Relative Clause Attachment. *Proceedings of the Speech Prosody Conference*. Dublin, Ireland.
- [6] Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders* 34(2), 163–175.

Emotion, reference, and perspective-taking in children's language processing

Craig G. Chambers¹, Jared Berman², Melanie Khu², Justine Thacker², & Susan A. Graham²
¹University of Toronto, ²University of Calgary

A hallmark of socially-situated language use is a person's ability to consider the knowledge and perspective of other conversational partners. For example, a listener needs to track and integrate knowledge about a speaker's knowledge state to successfully identify linguistic intent. To date, much of the work in real-time comprehension has used sensitivity to a speaker's *visual perspective* as a measure of broader perspective-taking abilities. Here we describe work instead exploring children's *emotional* perspective-taking in communication. This is motivated by (i) the evolutionarily-entrenched status of emotion, (ii) the early emergence of emotional sensitivities in childhood, and (iii) the fact that emotion cues are communicated paralinguistically in speech, making these cues part of the "bottom-up" signal rather than contextual information separate from speech. In our initial eye tracking studies investigating emotion and linguistic reference, 3-, 4-, and 5-year-olds heard instructions of the type *Point to the doll* where the NP was indeterminate but could be disambiguated by happy- or sad-sounding intonation (e.g., the display contained an intact and broken doll, among other objects). The results showed children's ability to rapidly identify emotion in speech and use this information to infer referential intent, as well as developmental gains over the age range studied. However, we also found intriguing disconnects between gaze and overt behavior (pointing), as well as differences related to emotional valence (happy vs. sad intonation). We also found that real-time use of emotional cues was modulated by knowledge about the talker: talkers who had unusual emotional reactions to objects (on early trials) made child listeners less likely to disambiguate reference using emotional intonation.

Our most stringent test of perspective-taking involved a competitive game task with an adult confederate. Here, a "win" for the child participant on a game round corresponded to a "loss" for the adult partner, and vice versa. On every trial, both players viewed a visual display containing two doors: the child's door and the confederate's door (each player had their own display screen). The players knew there would be a sticker hidden behind only one door and that the location of the sticker would determine which player would win a real-world sticker. In addition, the adult player always received advance knowledge of (but no control over) the sticker's location on each trial (doors opened first on the adult's screen). As the adult player learned the outcome, she produced a referentially ambiguous sentence (e.g., *Look! There it is*), spoken with either happy or sad emotional intonation. Gaze patterns showed children bypassed egocentric interpretations and anticipated their own door in the sad-speech condition and the adult's door in the happy-speech condition. These sensitivities, however, were less apparent in pointing decisions, once again highlighting a gap between children's implicit sensitivities and overt reasoning. We next correlated gaze patterns with a range of individual differences measures. We found that on-line linguistic perspective-taking was positively correlated with receptive vocabulary and a measure of (nonlinguistic) emotional perspective-taking. In contrast, there was no correlation with measures of individual differences in *visual* perspective-taking, conflict or delay inhibitory control, or working memory. Thus the abilities supporting communicative perspective-taking in cases involving *emotion* are not the same as those involving shared/privileged visual access to display objects. This final outcome is inconsistent with any claim that common ground in language processing involves a type of "general-purpose" mechanism for tracking partners' knowledge states and beliefs.



Example of visual display on core perspective-taking task, showing the timing between display status, speech and the measurement of anticipatory gaze fixation. The child's door is blue and the adult partner's door is red. A happy voice should trigger anticipatory fixations to the adult speaker's door, and a sad voice should trigger fixations to the child listener's door.

Do posh ducks say qu[ɑ:]ck?: Investigating the cognitive representation of dialect variation

Martha Austen | *The Ohio State University*

During speech comprehension, listeners decode both linguistic (*what* the talker says) and social (*who* the talker is) information. Listeners' social evaluation of a talker can be influenced by fine-grained phonetic information in the speech signal (e.g. Plichta & Preston 2011). Similarly, listeners' linguistic perception of the speech signal—e.g. what phoneme they hear—can be influenced by social information about a talker (e.g. Strand & Johnson 1996). This suggests that listeners draw on stored links between phonetic detail, linguistic categories, and social categories when perceiving speech. However, it is unclear to what degree these stored links are shared by social and linguistic perception: is the knowledge a listener uses to socially evaluate a talker the same knowledge they use to understand what that talker is saying? This study aims to disambiguate models that assume the answer is yes (e.g. Kleinschmidt et al. 2018) from those that imply separate sets of representations (Labov 2006; Eckert and Labov 2017) by looking at patterns of overgeneralization in cross-dialectal speech perception. It asks whether individual American listeners will overgeneralize in the same way across their social and linguistic perceptions of [ɑ:] in southern British English (SBE), arguing that mismatches would occur only under a model with separate sets of representations.

SBE has two phones—short [a] for TRAP words (*trap, cat, hat*, etc.) and [ɑ:] for BATH words (*bath, ask, path*, etc.)—where general American English only has one, a short [æ] (perceptually similar to [a]); Americans may therefore be unaware of which words belong to the BATH set in SBE. The TRAP/BATH distinction is phonetically conditioned, with some lexical exceptions in the TRAP set such as *g[a:]s* (referred to here as GAS words). While some Americans may have accurate representations of which words SBE speakers produce with [ɑ:] (“accurate Americans”), others may overgeneralize: either (1) across their entire /æ/ phoneme, believing that SBE speakers also produce TRAP/GAS words with [ɑ:] (“phonemic overgeneralizers”), or (2) across /æ/ in the relevant phonological environments without learning the lexical exceptions, believing that SBE speakers produce GAS but not TRAP words with [ɑ:] (“allophonic overgeneralizers”). In social perception, [ɑ:] should be linked to categories like “posh” and “British” in BATH words (Gupta 2005), with overgeneralizers extending these meanings to GAS/TRAP words. Under a model with shared representations for linguistic and social perception, individuals should overgeneralize to the same extent across both types of perception.

To test this prediction, approximately 90 Americans and 30 SBE controls will be recruited from the crowdsourcing website Prolific Academic. To measure linguistic perception, participants will complete a lexical decision task with TRAP, BATH, and GAS words produced with [ɑ:], e.g. *tr[ɑ:]p*, *b[ɑ:]th*, and *g[ɑ:]s*. SBE controls and 'accurate Americans' are expected to accept only BATH class stimuli as words, whereas overgeneralizers should also accept GAS and, for phonemic overgeneralizers, BATH words (see Table 1). To measure social perception, participants will complete a modified matched guise task in which they hear two versions of the same TRAP, BATH, or GAS sentence, one with [a] and the other with [ɑ:] (e.g. *I took a b[a]th* vs *I took a b[ɑ:]th*), and are asked to choose which version sounds more X, e.g. more posh. All participants should hear [ɑ:] as more posh than [a] in BATH sentences, but for TRAP and GAS sentences, SBE controls and accurate Americans are expected to hear [a] as more posh (with tokens like *g[ɑ:]s* sounding like someone trying—and failing—to sound posh). Allophonic overgeneralizers should hear [ɑ:] as more posh in BATH/GAS sentences but not TRAP sentences, and phonemic overgeneralizers should hear [ɑ:] as more posh in all three sentence classes (Table 2). Crucially, individual participants' performances across the two tasks will be compared.

Preliminary results suggest that in linguistic perception, most Americans are allophonic overgeneralizers but display little or no generalization to TRAP words. However, in social perception some Americans *do* overgeneralize to TRAP words, hearing productions like *tr[ɑ:]p* as “posh” and “authentically British”, versus *tr[a]p* (the actual SBE pronunciation) as American-influenced. This mismatch between linguistic and social perception suggests that social perception may tend to occur on a more abstract level of representation, and that models of speech perception cannot take for granted that listeners use a single set of links to understand both what a person says and who that person is.

		Predicted Response in Lexical Decision Task if Representation of [ɑ:] is ...			
stimulus type	example	No Representation (No adaptation)	Phoneme (phonemic over-generalization)	Phoneme + Conditioning Environment (allophonic over-generalization)	Accurate (either word-specific or phoneme + conditioning environment + lexical exceptions)
BATH with [ɑ:]	b[ɑ:]th (SBE - b[ɑ:]th)	no	yes	yes	yes
TRAP with [ɑ:]	tr[ɑ:]p (SBE- tr[a]p)	no	yes	no	no
lexical exception with [ɑ:]	g[ɑ:]s (SBE- g[a]s)	no	yes	yes	no

Table 1: Predicted responses on the lexical decision task by representation type.

		[ɑ:] sounds _____ than [a] if representation is ...			
stimulus type	example	No Representation (No social meaning)	Phoneme (phonemic over-generalization)	Phoneme + Conditioning Environment (allophonic over-generalization)	Accurate (either word-specific or phoneme + conditioning environment + lexical exceptions)
BATH with [ɑ:]	b[ɑ:]th (vs b[a]th)	no difference	more posh	more posh	more posh
TRAP with [ɑ:]	tr[ɑ:]p (vs tr[a]p)	no difference	more posh	less posh	less posh
lexical exception with [ɑ:]	g[ɑ:]s (vs g[a]s)	no difference	more posh	more posh	less posh

Table 2: Predicted responses on the matched guise task by representation type.

References

- Eckert, Penelope, and William Labov. 2017. "Phonetics, phonology and social meaning." *Journal of Sociolinguistics* 21, no. 4: 467-496.
- Gupta, Anthea Fraser. 2005. "Baths and Becks." *English Today* 21, no. 1: 21–27.
- Kleinschmidt, Dave F., Kodi Weatherholtz, and T. Florian Jaeger. 2018. "Sociolinguistic Perception as Inference Under Uncertainty." *Topics in Cognitive Science*.
- Labov, William. 2006. "A Sociolinguistic Perspective on Sociophonetic Research." *Journal of Phonetics* 34 (4): 500–515.
- Plichta, Bartłomiej, and Dennis R. Preston. "The /ay/s Have It: The Perception of /ay/as a North-South Stereotype in United States English." *Acta Linguistica Hafniensia* 37, no. 1 (2005): 107–130.
- Strand, Elizabeth, & Keith Johnson. 1996. Gradient and Visual Speaker Normalization in the Perception of Fricatives. In *KONVENS*: 14-26.

Automatic vs. RT-Modulated Phonetic Convergence over Timescales and Contexts

Stephen J. Tobin, Marc A. Hullebus & Adamantios I. Gafos

Universität Potsdam

Reports of phonetic convergence have typically been on long timescales [1,2] and have relied on comparisons between baseline and test means of phonetic parameters values [2,3]. In the present investigation, we sought immediate effects of convergence within individual perception-action loops in a cue-distractor paradigm. We propose that convergence is automatic on this fast timescale but that it can be modulated/reversed by social factors on longer timescales [4,5,6].

In a cue-distractor paradigm, participants utter syllables in response to visual cues (“if you see ** say ka, if you see ## say ta”). Once the visual cue appears, but before the participant produces their response, participants hear a distractor. The task is conducted in a laboratory setting. Previous applications of this paradigm have revealed perceptuomotor integration—fine-grained variation in response times (RTs) according to the degree of featural match (voicing, place of articulation) between cue and distractor [7,8,9].

In our extension of the paradigm, we controlled, for the first time, the phonetic distance between individual participant mean VOT and distractor VOT. We did this by registering participant-specific VOTs in a baseline block of 100 trials (50 /ta/, 50 /ka/) without any distractor. In the experimental block (N=720) every trial contained a distractor (360 /ta/, 360 /ka/; 50% matching cue+distractor, 50% mismatching). /ta/ and /ka/ distractor stimuli were drawn from 9-step VOT continua (45-85 ms) created by resynthesis in Praat.

In order to detect subtle effects of convergence (cf. [2,3]) and to adequately represent the predictor variable of phonetic distance, we derived two new variables. First, to represent phonetic distance between participant baseline and distractor tokens, we subtracted the mean of each participant’s baseline VOT (VOT^{mb}) from each trial’s distractor VOT (δ^{db} , distractor-baseline differential). On this scale, zero means that $VOT^{mb} = VOT_{distractor}$. Positive values indicate distractors above participant VOT^{mb} and negative values indicate distractors below participant VOT^{mb} . Next, to detect convergence among response VOTs in spite of variation in syllable (σ) duration, we normalized response VOTs by σ duration. σ duration covaries with VOT [10] and could confound subtle effects of convergence. Thus, we divided response VOTs by σ duration (VOT/σ). As with δ^{db} , we converted this quotient into a difference score ($[VOT/\sigma]_{response} - [VOT/\sigma]_{baseline}$) on a participant-wise basis (δ^{rb} , response-baseline differential). On this scale, zero means that $[VOT/\sigma]_{response} = [VOT/\sigma]_{baseline}$. Positive values indicate normalized response VOT quotients above baseline quotients and negative values indicate they were below baseline. Convergence is indicated by positive slopes between δ^{db} and δ^{rb} .

We present the results of 22 participants. Linear mixed effects regression in R [11] revealed significant effects of RT, and significant interactions of $\delta^{db} \times Match$, $RT \times Match$, and $\delta^{db} \times RT \times Match$. Figure 1 depicts regression lines from representative points along our continuous predictor of RT (left to right) in /ta/ (top row) and /ka/ (bottom row) Cue Syllable conditions. We highlight our two main results. First, there is an effect of δ^{db} (x-axis) on δ^{rb} (y-axis)—the positive slopes indicate convergence. Second, there is a modulating interaction of RT and Match—longer RTs yield less convergence (flatter slopes) and mismatching cue+distractor pairs yield less convergence than matching ones (flatter slopes in Mismatch condition).

Our results constitute the first demonstration of trial-to-trial phonetic convergence, controlling for phonetic distance. We will discuss how our results fit within the range of convergence [6], maintenance [12,13], and divergence [5] effects that have been reported in other work and on other timescales [1]. We specifically propose that in spite of automaticity of convergence at the fastest timescales, variability in interactional partners [13], social and sociolinguistic linguistic factors [5,14] modulate accommodation at longer timescales.

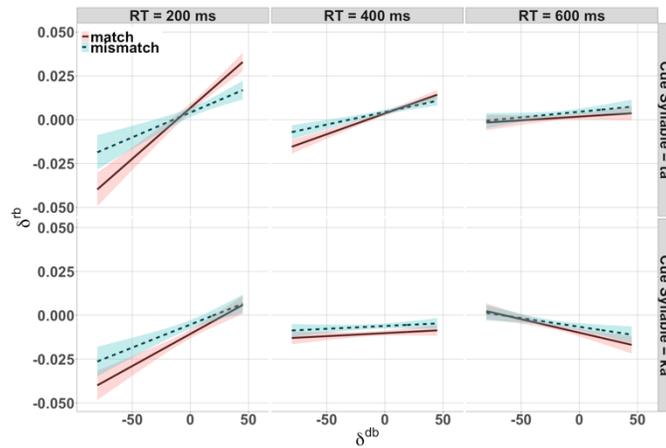


Figure 1: $\delta^{\text{db}} \times \delta^{\text{tb}}$ regressions with 90% CIs. Three representative points from continuous RT are shown from left to right. Match vs. mismatch cue-distractor conditions are shown in the top vs. bottom row, respectively. Positive slopes indicate convergence.

- [1] Major, R. C. (1992). Losing English as a first language. *MLJ*, 76(2), 190–208.
- [2] Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *JPhon*, 25(4), 421–436.
- [3] Tobin, S., Nam, H. & Fowler, C. (2017). Phonetic drift in Spanish-English bilinguals: Experiment and a self-organizing model. *JPhon*, 65, 45–59.
- [4] Krauss, R. M., Pardo, J. S. (2006). Speaker Perception and Social Behavior: Bridging Social Psychology and Speech Science. In P.A.M. van Lange (Ed.), *Bridging Social Psychology: Benefits of Transdisciplinary Approaches*, pp. 273–278. Lawrence Erlbaum Associates, NJ.
- [5] Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, N. Coupland & J. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). Cambridge: Cambridge University Press.
- [6] Babel, M. E. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *JPhon*, 40(1), 177–189.
- [7] Kerzel D. & Bekkering, H. (2000). Motor activation from visible speech: Evidence from stimulus response compatibility, *JEP:HPP* 26(2), 634–647.
- [8] Gallantucci, B., Fowler, C., & Goldstein, L. (2009). Perceptuomotor compatibility effects in speech, *APP* 71(5), 1138–1149.
- [9] Roon, K. & Gafos, A. (2015). Perceptuo-motor effects of cue-distractor compatibility in speech: beyond phonemic identity, *PBR* 22(1), 242–250.
- [10] Allen, S., Miller, J. & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *JASA* 113(1), 544–552.
- [11] Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *JML*, 59, 390–412.
- [12] Mitterer, H. (2014). Agreeing to disagree: Constant non-alignment of speech gestures in dialogue, *Journal of The Malta Chamber of Scientists*, 2(1), 8–14.
- [13] Sonderegger, M., Bane, M. & Graff, P. (2017). The medium-term dynamics of accents on reality television. *Language*, 93(3), 598–640.
- [14] Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *LabPhon*, 2(1), 125–156.

Processing of case variation in German prepositional phrases

Alexandra Engel¹, Adriana Hanulíková^{1,2}

¹ University of Konstanz, Germany

² University of Freiburg, Germany

Previous research has shown that speaker identity and speaking style affect the processing of morphosyntactic violations. That is, a grammatical error produced by a non-native speaker is not processed like one produced by a native speaker (Hanulíková et al., 2012). Similar effects have been shown by contrasting errors produced in casual and careful speech (Viebahn et al., 2017). However, it is not yet clear to which kind of errors language users adapt their processing. The present study examined how grammatical variation in prepositional phrases affects processing of written (using self-paced reading, SPR) and spoken language (using sentence repetition and completion task in careful and casual speaking styles).

In German, prepositions govern the case of the following noun. Corpus studies show that some prepositions alter between genitive and dative case without a semantic motivation (e.g., *wegen des Regens* vs. *wegen dem Regen*, “because of the rain”; cf. Di Meola, 2000). The combination of the genitival prepositions *wegen* (“because of”) or *während* (“during”) with the dative is considered non-standard and occurs more often in spontaneous speech, colloquial language use, and in dialects (cf. Scott, 2011, pp. 65-66). Hence, the question is how processing of standard and nonstandard variants differs and whether this is modulated by speaking style. To answer these questions, an SPR experiment and a sentence repetition and completion task with two speakers differing in speaking style have been set up.

The SPR experiment examined the processing of grammatical variation in combination with the prepositions *wegen* and *während*. We expected a processing benefit for the standard variants (cf. Squires, 2016). In contrast to this prediction, reading times of the participants ($n = 32$, mean age = 23.8) were faster in the noun region for non-standard dative variants compared to standard genitive variants. To examine whether the speaking style modulates expectations about case variants, we conducted a sentence repetition and completion experiment. After a familiarization phase with two talkers (one with a careful pronunciation and one with a casual pronunciation), participants ($n = 20$, mean age = 22.5) were required to repeat sentences produced by these two talkers. Importantly, critical case markings were masked by signal distortions. The results showed a main effect of speaker, suggesting that the processing of grammatical variants depends on speaking style-driven expectations. Furthermore, a main effect of block of each speaker and an interaction with the order of speaker presentation indicate that participants adjusted their responses throughout the experiment. That is, participants tended to use the standard genitive variant more frequently in the second block and particularly when the carefully pronouncing speaker was presented first. These findings have implications for the study of processing morphosyntactic variation and for the study of expectation-based language processing.

References:

- Di Meola, C. (2000). Die Grammatikalisierung deutscher Präpositionen. Tübingen: Stauffenburg-Verl. (Studien zur deutschen Grammatik, 62).
- Hanulíková, A.; van Alphen, P. M.; van Goch, M. M.; Weber, A. (2012). When One Person's Mistake Is Another's Standard Usage: The Effect of Foreign Accent on Syntactic Processing. In *Journal of Cognitive Neuroscience* 24/4, 878–887.
- Scott, A. K. (2011). Everyday Language in the Spotlight: The Decline of the Genitive Case. In *German as a foreign language journal* 1, 53–70.
- Squires, L. (2016). Processing Grammatical Differences. Perceiving versus Noticing. In Anna M. Babel (Ed.), *Awareness and Control in Sociolinguistic Research* (pp. 80-103). Cambridge: Cambridge University Press.
- Viebahn, M. C.; Ernestus, M.; McQueen, J. M. (2017). Speaking Style Influences the Brain's Electrophysiological Response to Grammatical Errors in Speech Comprehension. In *Journal of Cognitive Neuroscience* 29/7, 1132–1146.

Talk abstracts – 5th September

Integrating listener and speaker characteristics into the Coordinated Interplay Account

Katja Münster & Pia Knoeferle (Humboldt-Universität zu Berlin)
Katja.muenster@hu-berlin.de

It is by now well established that real-time language processing mechanisms can be influenced not only by linguistic (e.g., Cooper, 1974) but also by visual contextual information such as depicted objects or action events (e.g., Knoeferle et al., 2005; Tanenhaus et al., 1995). This interplay between visual and linguistic information during real-time language processing has been specified in the Coordinated Interplay Account (CIA, Knoeferle & Crocker, 2006, 2007; Knoeferle et al., 2014).

In the CIA, language processing, the mediation of attention and the reconciliation of linguistic and extralinguistic information happens in 3 forward-dependent processing steps that can overlap and occur in parallel. Step *i* deals with incremental sentence interpretation: The comprehender interprets an incoming word ($word_i$), on the basis of the already-built sentence interpretation and linguistic constraints. This yields an updated sentence interpretation (int_i). Using int_i , linguistic and long-term knowledge as well as existing expectations, the comprehender updates his expectations ($anti_i$). Working memory (WM) maintains the interpretation, expectations, and scene representations. The unfolding interpretation then guides the comprehender's (visual) attention to the corresponding scene input or its representation in working memory (step *i'*), eliciting a referential and an anticipatory search. The referential search is based on new referring expressions in int_i , and the anticipatory search is based on the linguistic expectations in $anti_i$. A merger additionally combines the information from the newly attended scene with the scene representation in WM ($scene_{i-1}$). This merging process yields $scene_i$. In the final step (step *i''*) the scene representations are related to the interpretation via coindexing of nouns / verbs with objects / action representations. The scene representation can further change the interpretation and the expectations (Knoeferle & Crocker, 2007; Knoeferle et al., 2014; see Guerra & Knoeferle, 2014, 2017 for more details on co-indexing beyond referential relations).

We extend the CIA towards the sCIA by explicitly integrating comprehender characteristics and the comprehender's representation of speaker characteristics (Figure 1). The representations of the speaker are part of the scene representation and are tracked via indices to WM representations. Indices also track listener characteristics in WM representations. The sCIA features ProCom as an index: ProCom comprises characteristics of the comprehender, such as literacy (cf., Mishra et al., 2011). These can modulate language comprehension as reflected in (visual) attention. In addition, the extended scene representation (of objects, a speaker's emotional facial expression) or indices to the interpretation (e.g., of speaker voice characteristics, see Figure 1) can modulate language comprehension as reflected in (visual) attention or event-related brain responses. The expectations resulting from listener and speaker effects can be socially-based and graded. The sCIA accordingly extends the 'ant' parameter of the CIA (capturing expectations) to 'ants'. Ant_s captures comprehender expectations motivated by both the comprehender's characteristics and the context (e.g., object, action, and speaker representations). The sCIA models these expectations as graded, which we capture by means of probabilities ranging from 0 to 1. These changes permit us to capture the effects of listener (e.g., literacy, and age) and speaker (e.g., voice and emotional facial expressions) characteristics on real-time language comprehension as exemplified by

results from a subset of the studies in Mishra et al. (2011), Münster (2016), Carminati and Knoeferle (2013) and Van Berkum et al. (2008). Figure 1 exemplifies the sCIA using the results of Van Berkum et al. (2008). In this EEG study, adults hear a sentence spoken in a voice that violates expectations regarding the e.g., the age of the speaker. Van Berkum et al.'s (2008) results show an N400 effect for speaker voice-sentence incongruities, such as *Every evening I drink some wine before I go to sleep* spoken in a child's (vs. an adult's voice).

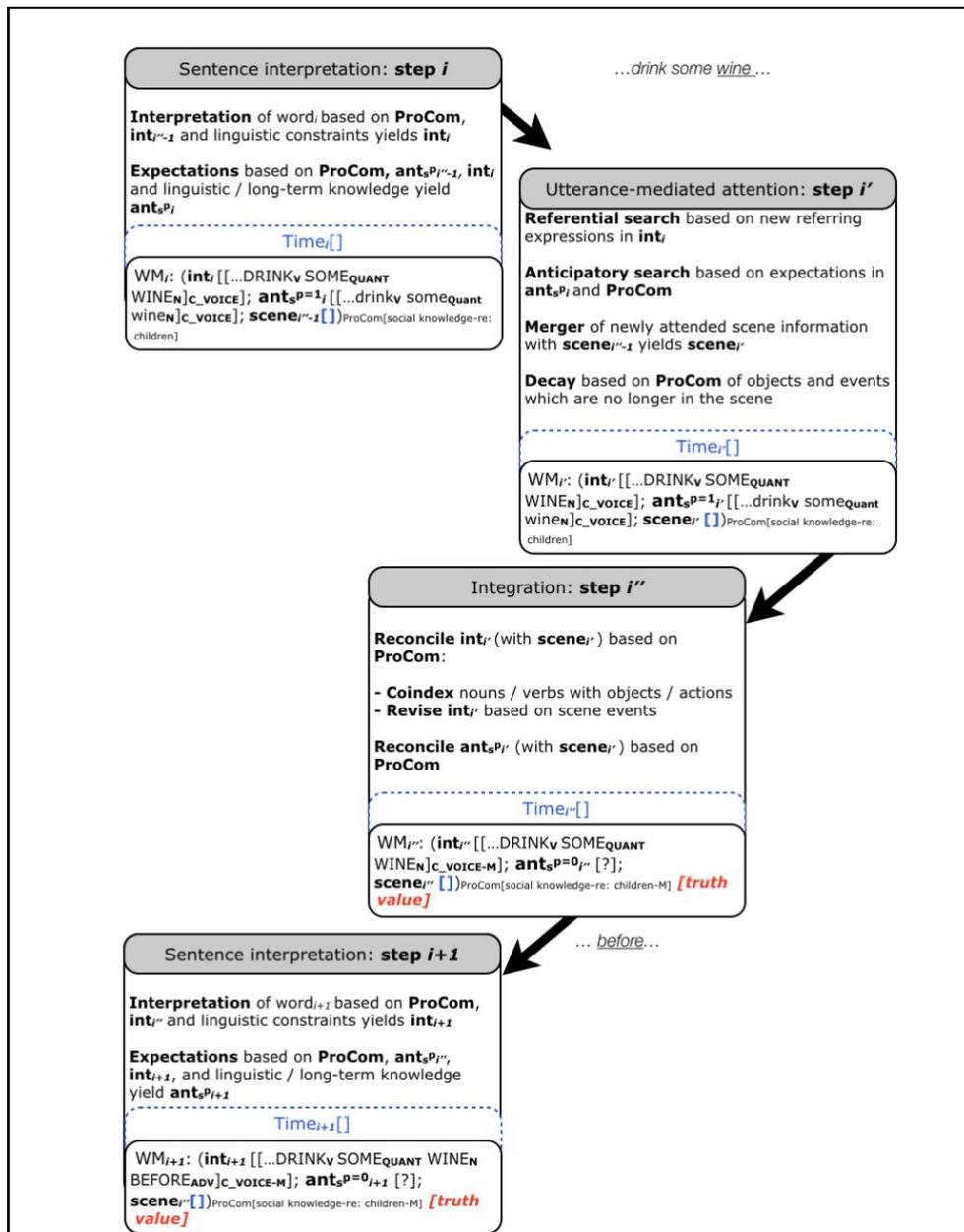


Figure 1: The social Coordinated Interplay Account (Münster & Knoeferle, 2018). ProCom can influence the interpretation (int) and (social-) context expectations (ant_s) that can influence sentence processing in real time. We exemplify the sCIA using the results by Van Berkum et al. (2008). In rendering the mental representation, we present the interpretation starting from the word *drink*. c_VOICE indicates that the comprehender hears a child's voice.

Integrating socially situated non-linguistic cues in pragmatic generalization

Amanda Pogue,¹ Sarah Brown-Schmidt,² & Chigusa Kurumada¹

¹University of Rochester, ²Vanderbilt University

Recent research demonstrates that listeners fine-tune their pragmatic expectations in response to evidence of under- or over-informativity, adapting their inferences in a talker-specific manner [1-3]. For instance, hearing a bare-noun expression (e.g., *Click on the bucket*) in the presence of two candidate objects (e.g., big and small buckets) causes a generalization of the inference that the talker may be under-informative in a future context. Such bare-noun expressions, however, can be sufficiently informative when accompanied by socially conveyed cues such as gaze, pointing, and affordances [4-7]. We test whether listeners can effectively integrate non-linguistic cues to modulate their pragmatic generalizations.

Exp(eriment)1 (n=101): P(articipant)s were randomly assigned to one of the three conditions: *Distance condition*, *Gaze condition*, *No Exposure condition*. At **Exposure**, in the *Distance* and *Gaze* conditions, Ps observed two talkers (T1, T2) instruct one another to arrange items on a board. Both talkers used bare-noun instructions (e.g., *Can you move the phone to the square on the left of the yarn?*). In the *Distance Condition*, for T1 the relative proximity of the objects disambiguated the referent (e.g., one of the objects was closer to, and therefore more accessible to, the listener), whereas for speaker T2 the candidate referents were equidistant (Fig. 1A-B). In the *Gaze condition*, T1's gaze direction indicated the target referent while T2's did not (Fig. 1C). At **Generalization**, Ps in all conditions viewed scenes containing images on a 2x2 image grid with one size-contrast pair and two unrelated singletons (Fig. 1E). Ps read a transcribed instruction that was either under-informative (color-modified, e.g., *Click on the green bottle*) or concisely-informative (scalar-modified, e.g., *Click on the wide bottle*) given the visual context, and selected a talker who was more likely to have produced the instruction. If a successful resolution of the bare-noun phrase in the exposure phase allows a listener to infer that the same talker would be more likely to provide informative instructions, Ps should select T1 for the scalar-modified (concisely-informative) instructions and T2 for color-modified (under-informative) instructions.

Results: The responses revealed a significant effect of exposure in *Distance condition*: T1, whose bare-noun instructions were visually disambiguated, was more strongly associated with the concisely-informative scalar instructions compared to the under-informative color modified instructions ($\beta = 8.16$, $z = 2.83$, $p < .01$, Fig. 2A top panel). No significant generalization was observed in the *Gaze Condition*.

In **Exp 2 (n=96)**: We repeated Exp. 1 with two critical changes: 1) In *Distance condition* we included both talkers faces (looking straight at the listener), which was not shown in Exp.1, to test if the competing facial cue might detract from the effect; and 2). In *Gaze + Point condition*, T1 both gazes and points at a referent (Fig. 1D).

Results: With the additional facial cue, the proximity manipulation in the *Distance condition* did not result in the predicted effect of scalar- vs. color-modified instructions on talker selections. The uninformative facial cue might have interfered and thereby weakened the effect of the proximity cue. In contrast, the *Gaze + Point condition* yielded the expected effect of scalar vs. color adjectives ($\beta = 1.16$, $p < .05$, Fig. 2A): Ps were more likely to select T1 and T2 for scalar and color instructions respectively. Overall, the results support the prediction that Ps incorporate multiple sources of non-linguistic information to assess likelihoods with which the talker informatively communicates their referential intentions.

Listeners thus assess talker's pragmatic reliability by learning from their use of language in context, facilitating generalization to unseen expressions. Yet, listeners did not indiscriminately use all possible cues, pointing to a system that goes beyond simple tracking of statistical co-occurrences between persons, context-cues, and linguistic expressions. Prior biases or learned utilities of social cues may shape pragmatic generalization.

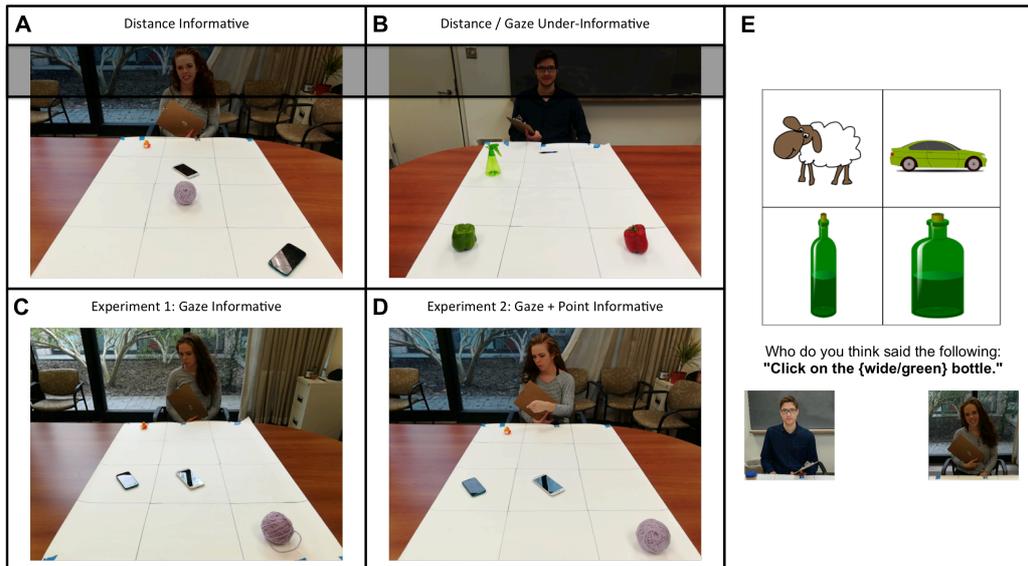


Figure 1: Example Exposure Phase trial in the Distance Informative Condition when the Context was Informative (target: phone in bottom right; A) or not (target: ambiguous; B), the shaded area was not visible in Experiment 1, but was visible in Experiment 2; example Informative Context Exposure Phase trial in the Gaze Informative Condition in Experiment 1 (target: phone on the left; C) and 2 (target: phone on the left; D) – the non-informative trials were identical across conditions (B). Example Generalization Phase trials (E) in Experiments 2 and 2; with an under-informative color-modified instruction (“green bottle”) and informative scalar-modified instruction (“wide bottle”).

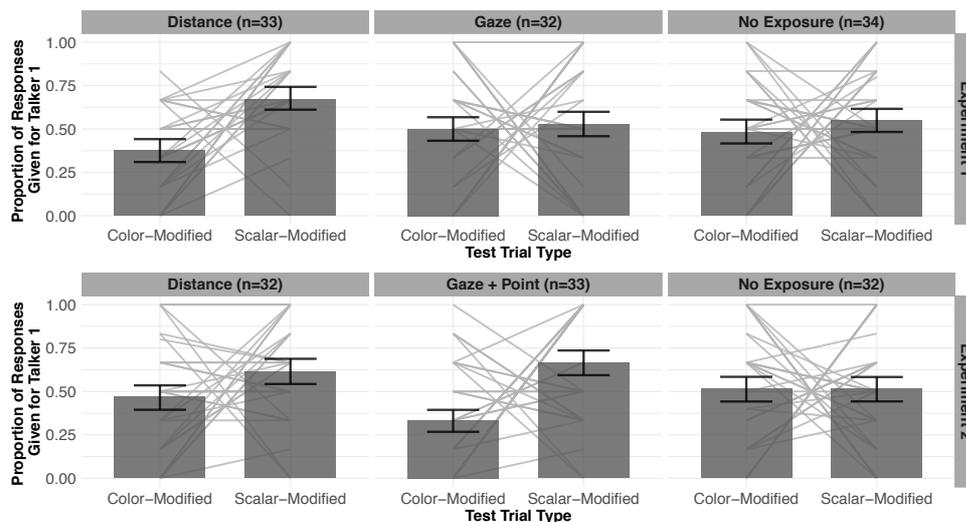


Figure 2: Proportions of Talker 1 selection (from the Exposure Phase) by Adjective Type in Test. The light gray lines = individual participant means. Error bars = 95% CIs.

[1] Grodner, D. J., & Sedivy, J. C. (2011). The effect of speaker-specific information on pragmatic inferences. [2] Pogue, A., Kurumada, C., & Tanenhaus, M. K. (2016). Talker-Specific generalization of pragmatic inferences based on under- and over-Informative pronominal adjective use. [3] Dix, S., Gardner, B., Lawrence, R., Morgan, C., Sullivan, A., & Kurumada, C. (under review). Integration of bottom-up and top-down information in online interpretations of scalar adjectives. [4] Brown-Schmidt, S., & Tanenhaus, M. K. (2008). Real-time investigation of referential domains in unscripted conversation: a targeted language game approach. [5] Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. [6] Hanna, J., & Brennan, S. (2007) Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. [7] Hanna, J., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: evidence from eye movements.

EFFECTS OF TALKER IDENTITY ON SPEECH COMPREHENSION ACROSS THE LIFESPAN

Adriana Hanulíková¹²³ (1University of Freiburg, 2FRIAS, 3University of Konstanz)
Adriana.Hanulikova@germanistik.uni-freiburg.de

Speech and faces provide a rich source of information about a talker identity. In a face-to-face conversation, listeners combine not only linguistic information but consider also information about who is speaking and how a message is communicated. Previous research suggests that perceived racial identity may alter comprehension, evaluation and perceived accentedness of speech from native and non-native speakers. Following expectation-based accounts, anticipating a talker's accent based on his/her ethnicity enhances intelligibility of speech if expectations are confirmed (e.g., Babel & Russell 2015; McGowan 2015). Bias-based accounts (e.g., Rubin, 1992; Kang & Rubin, 2009) such as reverse linguistic stereotyping predict comprehension difficulties due to stereotypical attributions to speaker's appearance. Such stereotyping however only occurs when native speech is perceived to be from non-native speakers.

In this paper, both proposals are examined by asking a) how perceived ethnicity shapes speech comprehension and accent ratings in foreign, regional, and native accents, and b) how effects of speaker identity on speech comprehension develop across the life span. Given that stereotypes and biases tend to be more difficult to suppress with increasing age, talker effects on speech comprehension may increase as well.

Teenagers (n = 78, mean age 14.1; range 12-15, native speakers of German), older adults (a preliminary group of n = 37; mean age 78; range 70-92, native speakers of German), as well as a middle age group (data collection ongoing) listened to nonsense utterances (e.g. "der alte Helm angelt" *the old helmet is fishing*) that were embedded in speech-shaped noise at 0 SRN. Utterances in Standard German, a regional (Palatinate) dialect, and a foreign (Korean) accent were presented over speakers along with photographs of white Caucasian or Asian women displayed on a computer monitor. Teenagers were asked to write down and older adults to repeat what the ostensible speaker said. Responses were scored based on % correct key word repetitions. Additional data on comprehensibility, accent ratings, working memory, and hearing ability (adults only) were collected.

In line with expectation-based accounts, speech intelligibility increased in case of a match between the talker's accent and perceived ethnicity of the ostensible speaker. Korean accent was more intelligible when presented along with the Asian guise, and Palatinate accent was more intelligible when presented along with the white Caucasian guise. Although speech comprehension in the standard condition was considerably higher than both accent conditions (with the Korean accent being more intelligible than the Palatinate accent), repetition accuracy did not depend on speaker ethnicity, in contrast to what would be predicted by both accounts. Older adults experienced greater difficulty understanding speech relative to children; however, talker effects were comparable across groups. Accent ratings revealed a main effect of ethnicity such that all varieties were rated as more accented in the Asian guise, in line with previous studies and the bias-based accounts.

Overall, the results show that social and linguistic information interact across the lifespan. Effects of talker ethnicity on comprehension however only occurred in noisy speech such as foreign or regional accented varieties. Both theories are partly confirmed, but seem to relate to different aspects of linguistic measures. Bias-based theories account well for metalinguistic and postperceptual processes (e.g. assessed by accentedness ratings), since accent ratings decreased for the Asian guise irrespective of the accent. Expectation-based theories account well for perceptual processes (e.g. assessed by intelligibility tasks), with accented speech being more intelligible if it is congruent with talker's anticipations driven by ethnicity.

References

- Rubin, Donald L. 1992. Nonlanguage factors affecting undergraduates' judgements of nonnative English-speaking teaching assistants. *Research in Higher Education* 33(4). 511-531.
- McGowan, Kevin. B. 2015. Social expectation improves speech perception in noise. *Language and Speech* 58(4). 502-521.
- Kang, Okim & Donald L. Rubin. 2009. Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology* 28. 441–456.
- Babel, Molly & Jamie Russell. 2015. Expectations and speech intelligibility. *Journal of Acoustical Society of America* 137(5). 2823-2833.

Novel lexical representations are shaped by speakers' in-group status and learners' in-group biases

S. Iacozza^{1,2}, A.S. Meyer^{1,3}, and S. Lev-Ari⁴

¹*Max Planck Institute for Psycholinguistics, Nijmegen*

²*IMPRS for Language sciences, Nijmegen*

³*Radboud University, Nijmegen*

⁴*Royal Holloway University of London, Egham*

Previous research has shown that individuals automatically encode others' social features [1] including whether or not they belong to their in-group [2]. Such social information has been reported to influence input processing. For example, people allocate extra attention towards in-group members, which leads to better and more individuated memories for in-group-related information than for information related to out-group members [3]. The goal of the current study is to investigate whether such memory patterns similarly influence the way people learn language from in- and out-group members.

In line with this prediction, previous research shows that people encode social features of speakers, such as their gender and age, along with the linguistic content they provide, and that this information influences on-line language processing. Indeed, not only do expectations driven by speakers' social identity modulate the ease of integrating linguistic meaning [e.g., 4] but they also influence how attentively the bottom-up signal is processed [5] and how language is interpreted [e.g., 6].

In the current study, we investigate the role of in-group status in a novel word learning paradigm. Specifically, we test 1) whether people encode the in-group status of speakers along with new words, and 2) whether in-group speakers lead to lexical representations that are more individuated compared to those elicited by out-group speakers and more so if learners exhibit greater personal in-group bias.

Participants performed a 2-session experiment. The goal of the first session was to implicitly teach participants which speakers belong to their in-group and which did not (i.e., from their own or a different university) by exposing them to facts about the speakers' habits. In the second session, participants indirectly learned novel words for novel objects from these supposedly in-group and out-group speakers via a comprehension task. Importantly, each target object was referred to twice: once by one speaker from each group. Following

the word learning task, participants performed a surprise source memory test for the words. Finally, participants' personal in-group bias was measured via a modified version of the perceptual matching task [8].

The results of the study confirmed our hypotheses: 1) Words were encoded together with information about the speakers. This was reflected by participants' greater tendency to confuse between speakers from the same university than between speakers of different university in the source memory task. 2) Participants' in-group bias modulated the differences in source memory between the two groups of speakers. Specifically, greater in-group bias predicted greater individuation of in-group members and lesser individualization of out-group members (see Fig.1).

Taken together, our findings show that social information such as in-group status, guides language learning and that personal biases modulate the magnitude of this effects. Thus, these findings highlight the importance of considering social factors when investigating language learning mechanisms and memory, and the importance of integrating the social context into models of learning and memory.

References

1. Bargh, J.A. et al. (2012). Automaticity in social-cognitive processes. *Trends in cognitive sciences* 16(12), 593-605.
2. Greenwald, A.G. and Banaji, M.R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychol. Rev.*, 102 (1), 4-27.
3. Hugenberg, K. et al. (2010). The categorization-individuation model: an integrative account of the other-race recognition deficit. *Psychol. Rev.*, 117(4), 1168.
4. Van Berkum, J. J. et al. (2008). The neural integration of speaker and message. *Journal of cognitive neuroscience*, 20(4), 580-591.
5. Lev-Ari, S. (2015). Comprehending non-native speakers: theory and evidence for adjustment in manner of processing. *Frontiers in psychology*, 5, 1546.
6. Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of language and social psychology*, 18(1), 62-85.
7. Ellenbogen, J. M. et al. (2007). Human relational memory requires time and sleep. *PNAS*, 104(18), 7723-7728.
8. Moradi, Z., Sui, J., Hewstone, M., and Humphreys, G. W. (2015). In-group modulation of perceptual matching. *Psychonomic bulletin and review*, 22(5), 1255-1277.

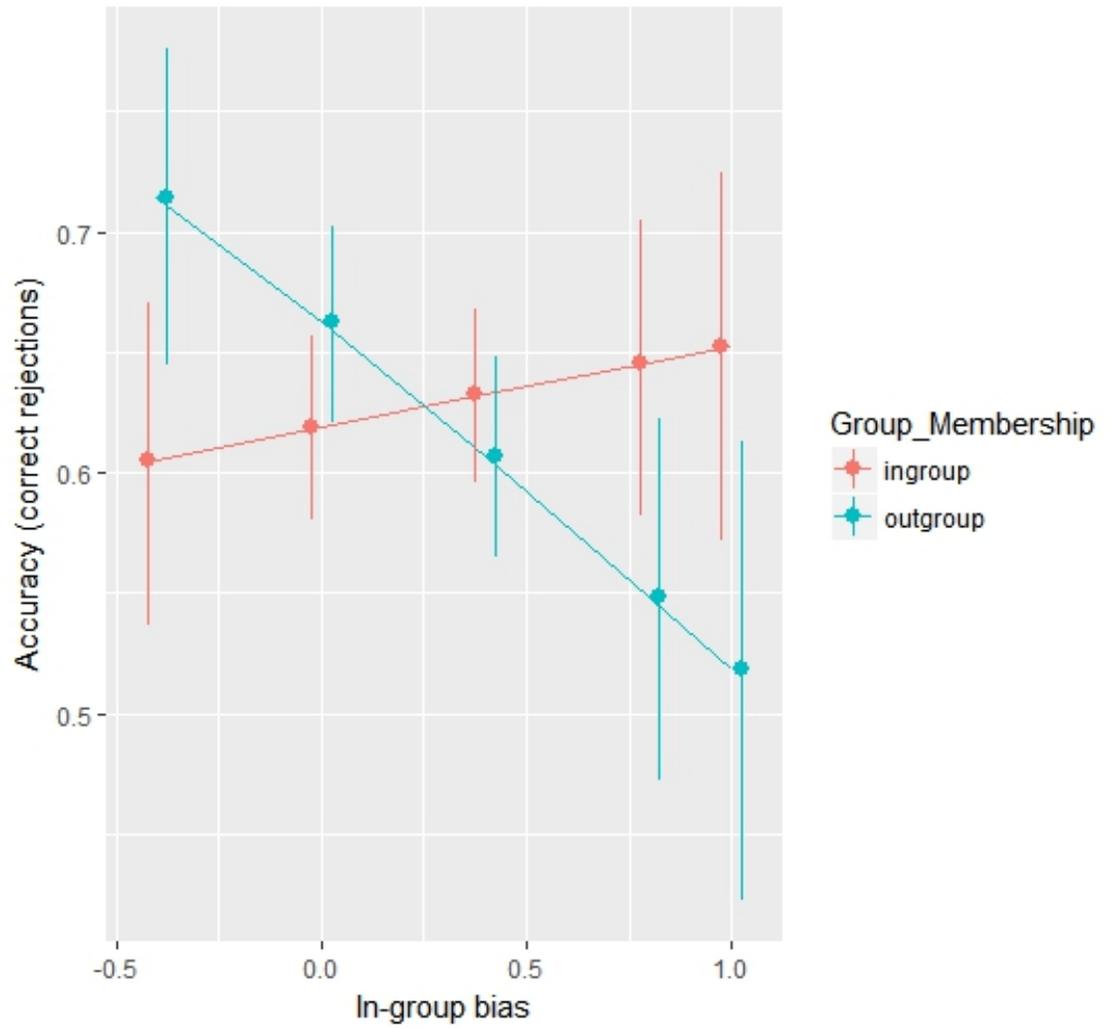


Figure 1: Correct rejections in the within-university trials as a function of speakers' in-group status and learners' in-group biases

Do these pants make me look fat? Influences of response delay on conversational meaning
Adriana Baltaretu, Craig G. Chambers, University of Toronto

In recent years there has been a resurgence of empirical work exploring aspects of conversational turn-taking as well as the timing of this process. Studies have suggested that the temporal gap between the contributions of dialogue partners (the 'inter-turn interval' [ITI]) not only reflects planning effort on the part of a speaker, but can also provide relevant social cues [1]. For example, increased ITIs can alter the perceived credibility of a speaker [2, 3, 4]. One goal of the current study was to explore how ITIs influence the pragmatic interpretation of speaker-communicated meaning, and to better understand the fine-grained sensitivity thresholds involved in this process. Participants listened to a series of conversational recordings where Yes/No questions posed by one person were responded to by another person (henceforth: Adjacency Pairs [APs]). On (24) critical trials, APs involved 'loaded' questions such as "Have you ever cheated on a test?" followed by the pro-social response ("No", in this case). These cases are intriguing because long delays in replying seem to invite us to discount the linguistic response and infer the respondent is unwilling to give a truthful reply. We varied ITI duration between questions and replies across 4 steps (500, 1000, 1500, 2000 ms). Our participant raters used a 5-point scale to gauge how sincere/truthful the respondent sounded in each AP. Each rater encountered a given AP only once, but across raters each AP occurred equally often with all four ITIs. On (24) filler trials, APs involved 'non-loaded' questions ("Do you have a pen in your bag?", again with varying ITIs). Talker voices and genders varied across trials.

A second focus of the study was to explore how the effect of ITI on conversational meaning may depend in part on raters' own processing speed/efficiency. E.g., interpretations at one particular ITI duration might be perceived differently by a 'fast' vs. 'slow' language comprehender, whereby the 'slow' individual perceives the respondent as more sincere. This would occur if our perceptions depend on an internal model of norms for conversational pacing that is based on our own information-processing traits. To this end we used three different participant groups thought to vary in speed/efficiency of processing, namely younger (N=30, mean age: 21) and older (N=20; age: 65+) native speakers; and younger non-native speakers (N=30, mean age: 21). We reasoned that ITI sensitivity might shift due to age-related cognitive slowing [5] and the delays in processing one's non-dominant language [6]. We also evaluated individual differences within each of these groups via a separate reaction-time task intended to provide a measure of each participant raters' processing speed.

The data (assessed using mixed-effect models) showed a pattern whereby perceived insincerity steadily and significantly increased as ITI duration increased (see Fig. 1). In addition, baselines were significantly shifted according to participant group: Older adults and non-native speakers rated respondents as more sincere across all ITI levels compared to younger adults (Fig. 2). However, within participant groups, we did not find correlations between ITI-driven interpretations and scores on the particular reaction-time task used here (which we note was not especially tailored to conversational or paralinguistic sensitivities).

Taken together, the current results show that fine-grained differences in ITIs add important and graded cues for comprehenders' interpretation of language in conversational contexts, even when the ITI range under investigation spans only 1.5 seconds in its entirety. Further, the results suggest language users who are less efficient at language processing have a shifted response to ITI gradients, whereby overall longer intervals are necessary to infer that what a respondent says may in fact differ from the truth. Although additional research is needed to rule out different socially-grounded conventions for conversational timing in older adult and nonnative speaker populations, we suggest our data indicate that aspects of meaning computed from paralinguistic cues can be nuanced by one's own processing traits.

References:

- [1] Sporer, S.L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20*, 421-446.
- [2] Brennan, S.E. & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory & Language, 34*, 383-398.
- [3] Ozuru, Y. & Hirst, W. (2006). Surface features of utterances, credibility judgements, and memory. *Memory and Cognition, 34*(7), 1512 - 1526.
- [4] Buller, D.B., Burgoon, J.K., Buslig, A.L.S., & Roiger, J.F. (1994). Interpersonal deception VIII. Further analysis of nonverbal and verbal correlates of equivocation from the Bavelas et. al (1990) research. *Journal of Language and Social Psychology, 13*(4), 396-417.
- [5] Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403.
- [6] Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*(1), 1-25.

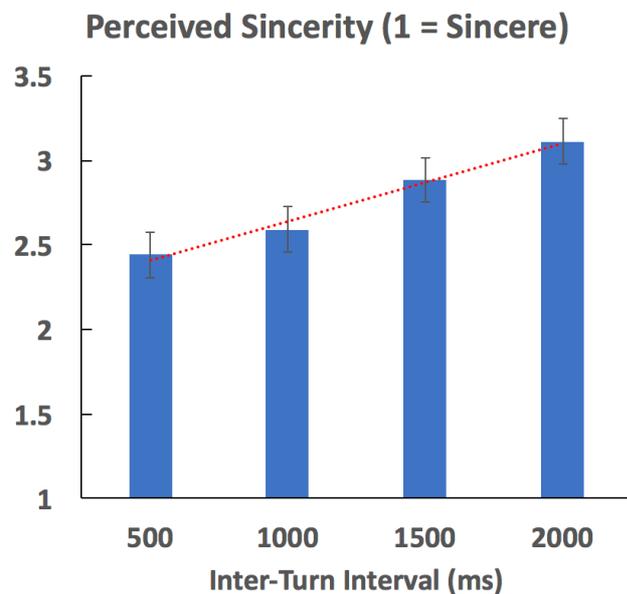


Figure 1. Differences in perceived (in)sincerity of respondent's reply across ITI steps, collapsed across all groups

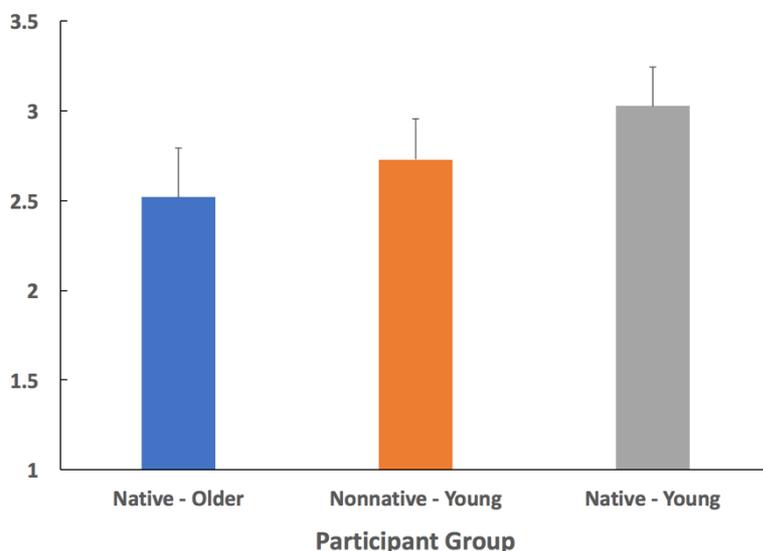


Figure 2. Differences in average perceived (in)sincerity by group

**VARIATION IN FRENCH PARTIAL INTERROGATIVES:
SOCIAL MEANING AS A KEY FACTOR TO UNDERSTAND SOCIOLINGUISTIC NORM VIOLATIONS**
- Gabriel THIBERGE & Barbara HEMFORTH (LLF, CNRS, Université Paris-Diderot, LabexEFL) -
Contact: gtac@tuta.io

French partial interrogatives show considerable variation (Coveney, 2011). The wh-element can be in a declarative argumental position (in situ, IS, “Ils mangent où ?”/“They eat where?”). It can also be in a fronted position, with subject-verb inversion, FINV (“Où mangent-ils ?”/ “Where eat they?”) or without inversion, F (“Où ils mangent ?”/“Where they eat?”). Syntactic constraints have been used to explain the differences in frequency of use for those variants: e.g. fronting is often analysed as structurally more ‘complex’ because of an underlying movement (Jakubowicz, 2011).

From a sociolinguistic view, previous research mostly studied the preferred use of one variant or another by specific social groups (Quillard, 2001). We extend this perspective in the framework of social meaning games (Burnett, 2017), where variation is a tool to socially position oneself in a specific context of interaction. To that end, we ran a series of three acceptability judgment tasks (AJT) and a written matched-guise test (MGT, Lambert, Hodgson, Gardner, & Fillenbaum, 1960).

Experiment 1: We ran an online written AJT (57 participants, aged 18-72, mean 28, median 33; 15 items; scale 1-10 based on French schooling system). It shows that FINV interrogatives are judged significantly better than IS questions (see Figure 1; $ps < .001$). This preference is entirely due to judgments of participants older than 30 years (interaction age*question type: $p < .01$; see Figure 2), which argues for a ‘generation gap’ among French speakers with regards to the consideration of all three interrogative variants.

Experiment 2: An online MGT provides a more fine-grained account: the perception of a French speaker's social persona (e.g. Eckert, 2008; Ochs, 1992) is affected by the syntactic structure he/she uses when producing partial interrogatives. 58 new participants from various social and educational backgrounds were presented with three interviews by a “journalist” that contained only one of the three question types. They judged the journalist on a variety of properties (Figure 3). Focusing on FINV vs. IS constructions, the use of FINV leads to a biased social perception on the receiver's part. People using FINV structures are thought to be of a higher social status, to have studied for a longer time, or to be older but also to be less relaxed.

Experiment 3: In an online AJT (44 new participants, 30 new items, scale 0-10), we integrated the questions in a formal or informal context. Results confirm the preference for FINV constructions over the two other variants as well as the age factor. Data also show that the preference for FINV constructions is context-dependant, in particular for participants 30 years and older. They judge FINV interrogatives significantly better ($p < 0.02$) in formal contexts than in informal ones (Figure 4).

Experiment 4: We replicated Experiment 3 with the exact same stimuli recorded in audio form (4 voices: 2 males, 2 females, in a balanced way across items). AJT was split in two questions, regarding both “well-formedness” and “suitability to the context”. Preliminary results (Figure 5) further confirm the age factor, but shed a new light on the context parameter: FINV questions are still judged as the ‘best French’ among all three structures across participants, but they are not judged as the ‘most adapted-to-the-context’. This confirms, experimentally, a bias in classic written AJT questionnaires regarding the well-formedness of French interrogatives. Moreover, while participants older than 30 no longer exhibit a preference among structures, participants under 30 seem to prefer IS structures.

Our AJT results, in combination with the MGT results, argue for the idea that French native speakers do make a meaningful choice when considering either one of the variants available to them for partial interrogatives. They attach social positioning values on various dimensions to each variant, and AJT ratings reflect these values. Whether that choice is conscious or not in production remains a question to be studied in future research.

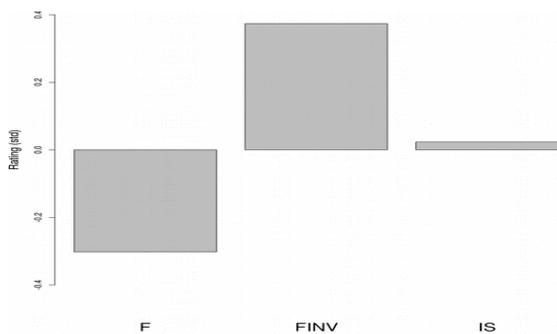


Figure 1: AJT 1 overall results (std.)

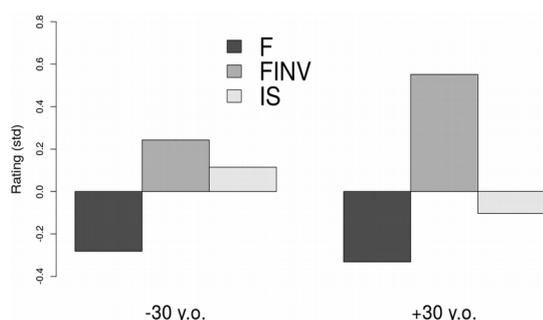


Figure 2: AJT 1 results by age (std.)

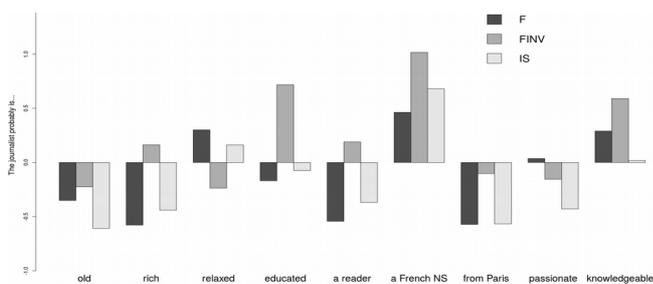


Figure 3: MGT global results (std.)

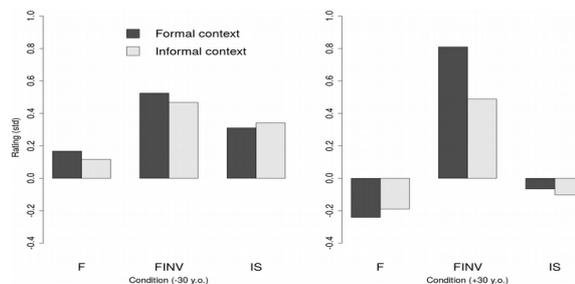


Figure 4: AJT 2 global results (std.)

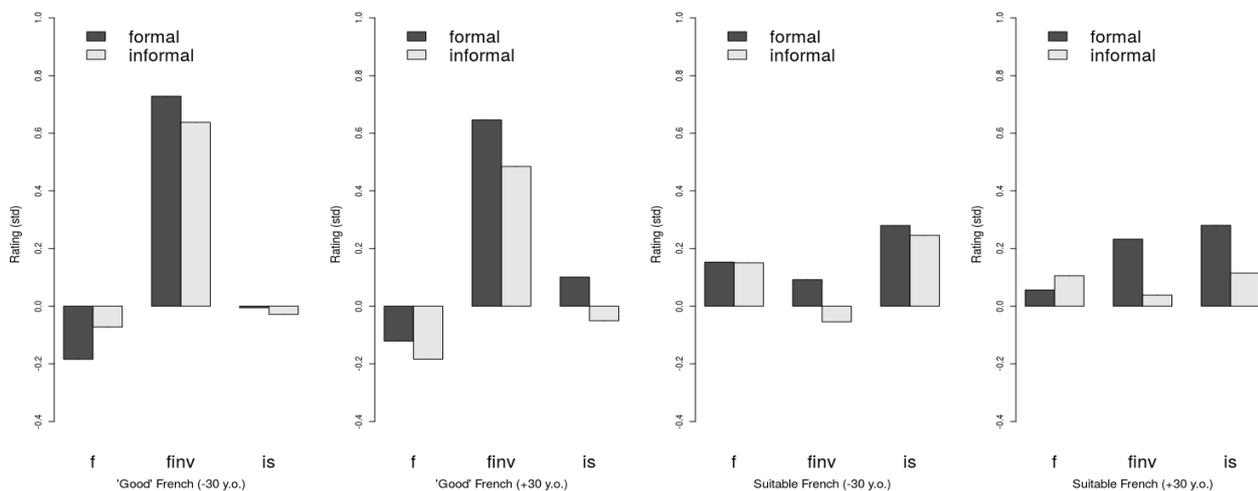


Figure 5: Audio AJT overall preliminary results (std.)

REFERENCES

- BURNETT, Heather (2017), Sociolinguistic Interaction and Identity Construction: The View from Game-Theoretic Pragmatics, *Journal of Sociolinguistics*, 22:1, pp. 238-271.
- COVENEY, Aidan (2011), L'interrogation directe, *Travaux de linguistique* 2011/2, n°63, pp. 112-145.
- ECKERT, Penelope (2008), Variation and the indexical field, *Journal of Sociolinguistics*, 12, pp. 453-476.
- JAKUBOWICZ, Célia (2011), Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French, in *Lingua*, 121, pp. 339-351.
- LAMBERT, Wallace E., Richard C. Hodgson, Robert C. Gardner & Samuel Fillenbaum (1960), Evaluational Reactions to Spoken Languages, *Journal of Abnormal and Social Psychology*, 20.1, pp. 44-51.
- OCHS, Elinor (1992), Indexing gender, in Alessandro Duranti and Charles Goodwin (eds.), *Rethinking Context: Language as an Interactive Phenomenon*, Cambridge, U.K.: Cambridge University Press, pp. 335-358.
- QUILLARD, Virginie (2001), La diversité des formes interrogatives : comment l'interpréter ?, *Langage et société*, vol. 95, no. 1, pp. 57-72.

Poster abstracts

REFERENCE RESOLUTION AND THE INTEGRATION OF REFERENTIAL VISUAL CUES

Mirjana Sekicki and Maria Staudte (Saarland University)

mirjana@coli.uni-saarland.de

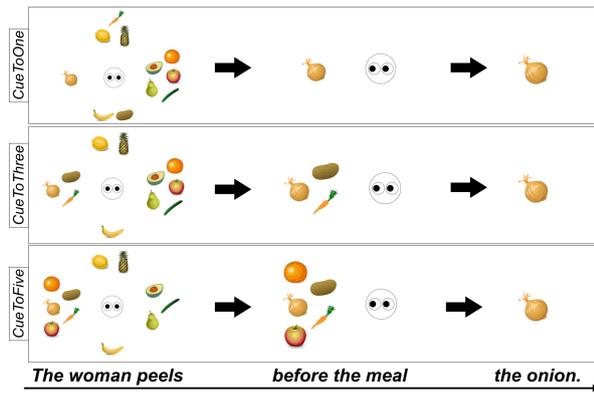
Visual cues are found to aid situated sentence processing by helping identify the target object. Specifically, referential speaker gaze, typically occurring prior to uttering the reference, has been shown to not only lead to shorter reaction times and higher accuracy on subsequent tasks, but also to a reduction of cognitive load induced by processing the subsequent reference. However, even though the cue can lead to a reduction of cognitive load on the referent noun, no indices of increased load were found on the cue itself, which would compensate for the reduction of load on the reference. Aiming to better understand the integration of visual cues in sentence processing, and the gradual reduction of referential uncertainty, we manipulated the specificity of the visual cue and the referent noun, as well as the point of reference resolution.

Two eye-tracking visual world experiments recruited 30 native German speakers each, and examined their visual attention together with the simultaneously induced cognitive load (measured with the Index of Cognitive Activity). Participants listened to simple sentences while viewing visual displays with a number of objects and a visual cue, in the form of schematic speaker gaze. In both experiments linguistic stimuli did not change within an item. Fig. 1a) illustrates that in Exp. 1 the visual cue reduced the referential uncertainty from eleven objects to one (*CueToOne*), three (*CueToThree*) or five (*CueToFive*). Subsequently, the referent noun named one object and completely disambiguated the target. The results show that even though the gaze cue directed participants' visual attention to the cued (group of) object(s), a graded difference in cognitive load was found only on the referent noun. Fig. 2 shows that *CueToOne* induced significantly lower load on the reference than the other two conditions ($p = .004$). *CueToThree* was positioned between *CueToOne* and *CueToFive*, but the differences did not reach significance.

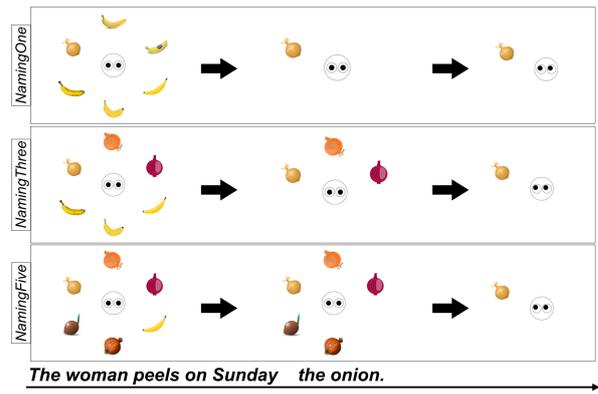
In Exp. 2, we swapped the roles of the visual cue and the reference, hence, the referent noun selected one (*NamingOne*), three (*NamingThree*) or five (*NamingFive*) objects in the scene, while gaze, now appearing after the noun, disambiguated the referent by cuing the specific target object (Fig. 1b). We combined the mentioned visual conditions with manipulating the existence of the cue (*no gaze vs. referent gaze*). The eye-movements show that upon hearing the noun, the target was inspected more in *NamingOne* than in the other two conditions ($p < .001$; see Fig. 4a). The gaze cue was followed, as revealed by the high inspection probabilities in all three visual conditions during the cue (Fig. 4b). Finally, we found no significant results in the analyses of cognitive load: Neither processing of the reference, nor of the visual cue induced statistically significant effects.

In sum, we found that both linguistic and visual cues reduced uncertainty and elicited shifts in visual attention accordingly. Different cognitive load results were found for the disambiguating noun (Exp. 1) versus the disambiguating visual cue (Exp. 2): While the cognitive load associated with the target was, as a result of the **preceding gaze cue**, proportional to gaze specificity (Exp. 1), processing the visual cue as a uniquely identifying reference **after** the nominal phrase (Exp. 2) did not influence immediate cognitive load. Importantly, high accuracy of the answers participants' gave on post-trial questions about both the sentence content and the visual scene clearly suggest that participants paid attention and mastered the tasks well.

We argue that the visual cue in Experiment 1 led to an expectation for a set of specific nouns. The smaller the set, the greater the expectation and the lower the associated cognitive load for the target noun. In contrast, the ambiguous noun in Experiment 2 did obviously not lead to an expectation for a set of visual cues so that the target cue would require lower cognitive load. We conclude that visual cues can reduce referential uncertainty for upcoming nouns but that they do not underly the same kinds of expectations (and resulting cognitive costs) as linguistic references.



a) Experiment 1



b) Experiment 2

Fig. 1: Steps of uncertainty reduction in the three visual conditions (zooming in on the relevant objects for illustration purposes).

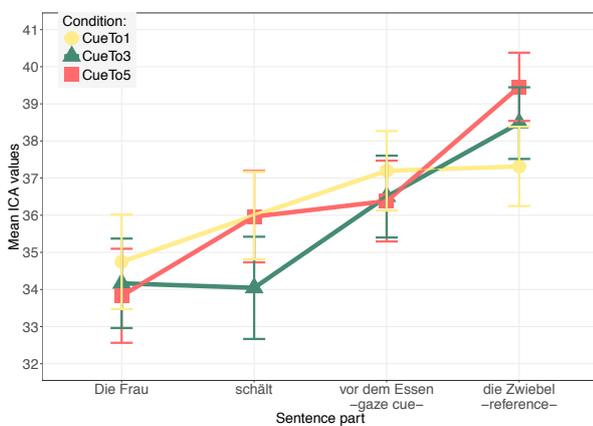


Fig. 2: (Exp. 1) Mean ICA values in four 600 ms time windows, illustrating the changes in cognitive effort in four segments of the sentence. Only the **final** segment shows sign. differences.

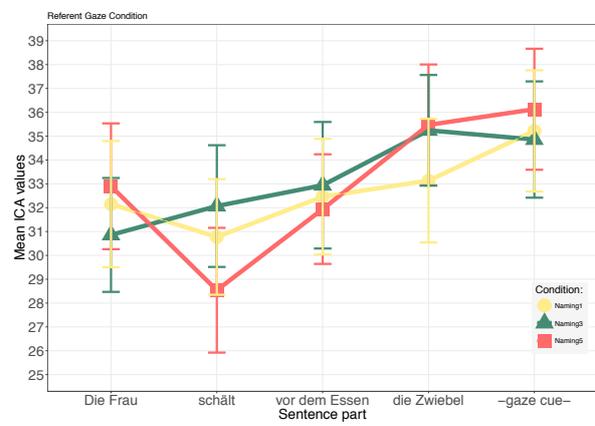
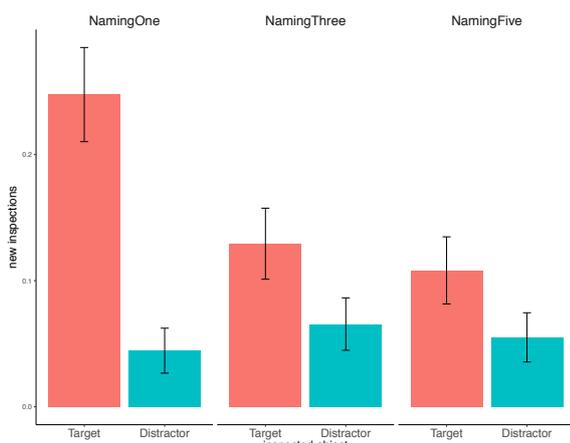
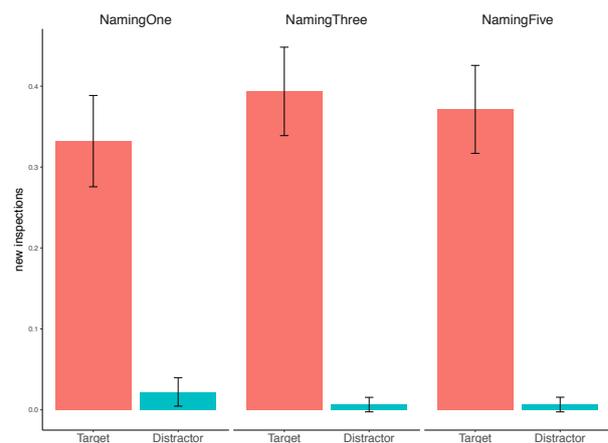


Fig. 3: (Exp. 2) Mean ICA values in five 600 ms time windows, illustrating the changes in cognitive effort in five segments of the sentence in the *referent gaze* condition. No sign. differences.



a) noun: Showing attention shift to the named objects rather than the distractor(s). Focus on several relevant objects reduces individual inspection time.



b) gaze: *referent gaze* condition. High inspection probability for target object shows gaze following.

Fig. 4: (Exp. 2) New inspections upon the noun (a) or gaze (b) directed towards the target object (red) and a representative distractor object (blue) in the three visual conditions.

**The difference is not orthographic, it is grammatical:
Situational variation of linguistic explanations**

Milena Kuehnast & Beate Lütke
Department of German Studies and Linguistics
Humboldt-Universität zu Berlin

The ability to tailor one's linguistic expressions according to salient features of the communicative situation starts developing already at the earliest stages of language acquisition. However, it has been shown not to follow a linear trajectory throughout childhood, adolescence and early adulthood (Berman & Nir-Sagiv, 2007, Kaplan & Berman, 2015). By investigating the linguistic features of written explanations, the present study explores how the acquisition of academic registers influences the intra-individual linguistic flexibility of German students enrolled in bachelor and master linguistics programmes.

The design of the present study incorporates individual properties and levels of academic education assumed to influence the development of linguistic flexibility operationalised as the ability to produce texts corresponding to the informational needs and the social status of the addressee.

We collected written explanations from German native speakers (N=70) divided into three groups based on their level of academic education in linguistics: beginners (1st and 2nd semester), intermediate bachelor students (3rd to 7th semester) and master students. They produced responses to a syntactic problem in two different communicative situations and completed the German version of the Big Five Inventory (personality structure), (Danner & al. 2016). The linguistic problems pertained to pairs of German sentences (1 & 2) depicting an ambiguity between a relative pronoun and a conjunction (das/dass 'that') or between a preposition and a conjunction (während: 'during/while').

- 1) Tim sieht, dass das Pferd einen neuen Sattel trägt.
'Tim sees that the horse wears a new saddle.'
- 2) Tim sieht das Pferd, das einen neuen Sattel trägt.
'Tim sees the horse that wears a new saddle.'

We varied the situation type by presenting the task as a) an e-mail from a 12-year-old pupil who requested help in explaining the orthographic differences between the two sentences, or b) as a part of a test at the end of a linguistics course, the task requiring an explanation of the structural differences between the sentences. Thus the linguistic context manipulated the representation of the communicative situation by altering the social and informational status of the addressee (a pupil or a lecturer).

The data were subjected to quantitative analyses assessing frequency and distribution of lexical, morphological and syntactic forms pertaining to academic or more informal discourses (Csomay, 2013; Czicza & al., 2012; Sardinha & Pinto, 2014). Qualitative analyses evaluated the differently framed explanations with respect to the type, the accuracy and the appropriateness of the arguments provided.

Preliminary results suggest a U-shaped trajectory in the development of linguistic flexibility during university education. Students at the intermediate level appeared to be less sensitive to the situational frame, frequently producing child-directed explanations similar to their lecturer-directed responses. In this group, some linguistic variables such as number of words or use of passive constructions showed correlations with personality traits such as extroversion and neuroticism, correspondingly. In contrast, first-year students and master students were more likely to adjust their explanations to the communicative needs of a school-age addressee, in terms of length, involvement markers, linguistic terminology and solutions.

The production pattern observed in our data is in line with investigations into the late stages of linguistic development providing evidence that the acquisition of conceptually written registers and highly specific forms of scientific discourses temporarily reduces the abilities of young adults to shift from academic to colloquial language (Berman & Nir-Sagiv 2007; Kaplan & Berman 2015).

References:

- Berman, R. A., & Nir-Sagiv, B. (2007). Comparing narrative and expository text construction across adolescence: A developmental paradox. *Discourse Processes*, 43(2), 79-120. doi:10.1080/01638530709336894
- Csomay, E. (2013). Lexical bundles in discourse structure: A corpus-based study of classroom discourse. *Applied Linguistics*, 34(3), 369-388. doi:10.1093/applin/ams045
- Czicza, D., Hennig, M., Emmrich, V., & Niemann, R. (2012). Zur Verortung von Texten zwischen den Polen maximaler und minimaler Wissenschaftlichkeit. Ein Operationalisierungsvorschlag. *Fachsprache*, 34(1-2), 2-44.
- Danner, D., Rammstedt, B., Bluemke, M., Treiber, L., Berres, S., Soto, C., & John, O. (2016). Die deutsche Version des Big Five Inventory 2 (BFI-2). *Zusammenstellung sozialwissenschaftlicher Items und Skalen*. Retrieved from doi:10.6102/zis247
- Kaplan, D., & Berman, R. A. (2015). Developing linguistic flexibility across the school years. *First Language*, 35(1), 27-53. doi:10.1177/0142723714566335
- Sardinha, T. B., & Pinto, M. V. (Eds.). (2014). *Multi-dimensional analysis, 25 years on. A tribute to Douglas Biber*. Amsterdam: John Benjamins.

Production/Interpretation Asymmetries in French Grammatical Gender

Célia Richy and Heather Burnett

Laboratoire de Linguistique Formelle (CNRS-Université Paris Diderot)

In this paper, we use methodology from sociolinguistics to contribute to an open question in the psycholinguistics of grammatical gender languages: *What is the relation between grammatical gender and social gender in the interpretation and production of human denoting noun phrases?* The paper focuses on French, where noun phrases with stereotypical masculine or feminine proper names/denotations are interpreted as referring to men or women, respectively (1a). Grammatically feminine nouns refer only to women (1b), but with most grammatically masculine human nouns (1c), the mapping between grammatical and social gender is less clear. Most traditional grammars and linguistics works (eg. Corbett 1991) propose that these so-called ‘generic masculines’ have a completely gender neutral interpretation; however, recent psycholinguistic research (Chatard et al. 2005, Brauer & Landry 2008, Gygax et al. 2008, 2012, Sato et al. 2013) has shown that when they are processed, they activate male mental representations creating a masculine bias in interpretation. In this paper, we present new evidence in favor of an **asymmetry** between production and interpretation with generic masculines: we show that although (as shown in other studies) the use of a generic masculine creates a bias for male interpretation, this bias does not appear to be reflected in language use.

Our study is inspired by research in sociolinguistics (McCauley & Brice 1997, Pabst et al. 2018) studying the distribution of references to men and women in constructed examples in English linguistics articles. Authors constructing linguistics examples are focused only on the form of their sentence, so this corpus provides a unique window into how unconscious gender biases can affect language production. Macaulay & Brice’s main quantitative results (replicated with more recent data by Pabst et al.) are that reference to men is far more frequent than reference to women across the board, but female referents are relatively more frequently in non-subject positions than male referents. They also found significant effects of author gender and publication date, with female authors and more recent publications referring more often to women. We replicated the M&B study using syntactic examples found in a major French linguistics journal from two periods: 1969-1971 and 2008-2017.

In **our first study**, we compared the distribution of noun phrases that have a clear masculine social gender interpretation (1a) with those with a feminine interpretation (1b) (N=2022). We found almost the same grammatical pattern as M&B: male reference is more frequent than female reference across the board (1654 occurrences vs 368), and we also found the same subject/non-subject asymmetry with female reference being more relatively likely in non-subject position than male reference (see Figure 1). We built mixed effects models with author gender, publication year and syntactic position as fixed effects and author as random effect. Unlike in the English studies, only syntactic position was significant ($p < 0.0001$). This result thus gives us an important generalization about social gender in French linguistics examples: subject position favors male reference compared to non-subject position. In **our second study**, we compared the syntactic distribution of true masculine DPs (1a), feminine DPs (1b) and generic masculines DPs (1c). Given all the psycholinguistic evidence showing that generic masculines are biased towards a male interpretation, if use of generic masculines mirrored their interpretation, we would expect these expressions to show the same general pattern as true masculines: there should be relatively more uses of generics in subject position than in non-subject position. However, as shown in Figure 2, generic masculines are equally likely in both subject and non-subject position. We suggest that this interpretation/production asymmetry is similar to the one observed for pronoun interpretation/production with implicit causality verbs (Stevenson et al. 1994, Rhode & Kehler 2014, Kehler & Rhode 2015, among

others), and we explore to what extent existing formal models for implicit causality bias (such as Kehler & Rhode 2015) can be extended to grammatical gender bias.

We have no preference regarding a talk or a poster.

Examples, Figures and References

- (1) a. Jean/Le père est arrivé. *Jean/The father arrived.* (male reference)
 b. Marie/La mère/la journaliste est arrivée.
Marie/The mother/the female journalist arrived. (female reference)
 c. Le journaliste est arrivé. *The journalist arrived.* (? reference)

Figure 1: Masculine and feminine references by syntactic position

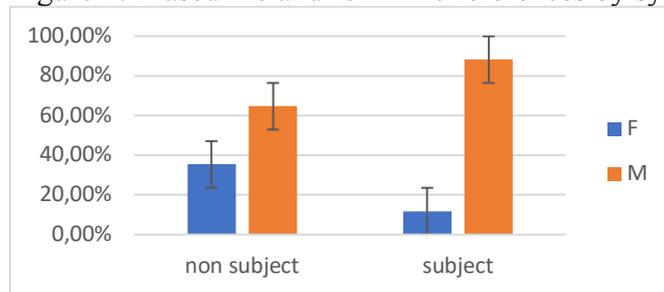
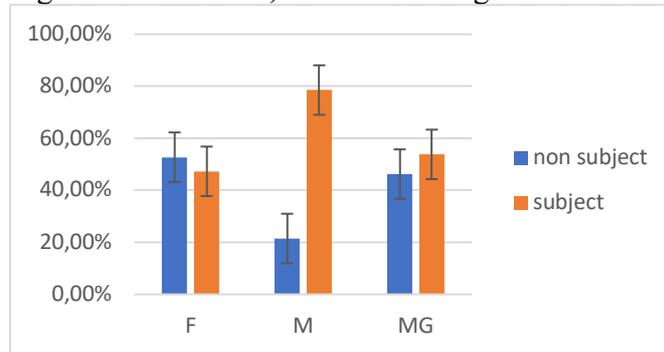


Figure 2: Masculine, feminine and 'generic masculine' by syntactic position



References

- Brauer, M., & Landry, M. (2008). Can a secretary of state become pregnant? The influence of the masculine generic on mental representations. *Année Psychologique*, 108(2), 243-272.
- Chatard, A., Guimont, S., and Martinot, D. (2005). Impact de la féminisation lexicale des professions sur l'auto-efficacité des élèves: une remise en cause de l'universalisme masculin? *L'année psychologique*, 105(2):249-272.
- Corbett, G. (1991). *Gender*. Cambridge University Press.
- Kehler, A & Rohde, H (2015). Pronominal reference and pragmatic enrichment: A Bayesian account. in *Proceedings of the 37th Annual Conference of the Cognitive Science Society*. pp. 1063.
- Gygax, P., et al. (2008). Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes*, 23, 464-485.
- Macaulay, M, and Brice, C. (1997). « Don't Touch My Projectile: Gender Bias and Stereotyping in Syntactic Examples ». in *Language*, Vol. 73, No. 4, pp. 798-825.
- Pabst, K. et al. (2018). Gender bias in linguistics textbooks. Paper presented at LSA 2018 annual meeting.
- Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, 29(8), 912-927.
- Stevenson, R., Crawley, R., & Kleinman, D. (1994). Thematic roles, focusing and the representation of events. *Language and Cognitive Processes*, 9, 519-548.

Toward an understanding of how nonbinary pronouns are learned and processed

Lauren Ackerman (Newcastle University), lauren.ackerman@ncl.ac.uk

There are few empirical studies examining how “singular *they*” is processed when paired with antecedents of ambiguous, unknown, or unambiguous gender (e.g. Foertsch & Gernsbacher 1997, Doherty & Conklin 2017). However, the primary focus of these studies was the use of singular *they* as a generic pronoun (e.g., 1a, 2a) rather than as a personal pronoun (1b, 2b), i.e., for a person of non-binary gender. This study aims to identify sources of variation in processing singular *they* with the goal of developing a theory about how stigmatised novel pronouns are (or are not) learned.

- 1) a. Anyone might be able to dress themselves every morning. (unbiased, generic)
- b. Taylor might be able to dress themselves every morning. (unbiased, specific)
- 2) a. A mechanic might be able to dress themselves every morning. (biased, generic)
- b. Joseph might be able to dress themselves every morning. (biased, specific)

The case for investigating singular (personal) *they* is particularly urgent, as the recent increase in visibility of nonbinary pronouns has led to outcries from even well-respected linguists that learning a novel pronoun is too difficult (Conrod 2017). Using the wrong pronoun for someone does lasting psychological harm for binary (and nonbinary) transgender people (McLemore 2015). Thus, determining how non-standard or novel pronouns are processed and learned does double duty by providing research that directly benefits a vulnerable population while also informing our understanding of how pronominal features are syntactically, pragmatically, and cognitively configured.

In this study, an eye-tracking while reading task examined reading behaviour in sentences like those in (1) and (2), with a particular focus on determining whether the specific (personal) use of singular *they* exhibited the established patterns of the generic. These established patterns were replicated, but did not extend to the specific condition (FF: $\chi^2(1)=5.1$, $p=0.025$). Demographic data was collected, including participant gender, age, level of education, introspection on use of singular *they*, LGBT affiliation, exposure to gender nonconformity in media and daily life, etc. This information was included in statistical models to ascertain the contribution of each factor to processing behaviours.

Unexpectedly, only two demographic effects approached significance. First, *exposure* is a coarse-grained measure of real-life experience with trans/nonbinary people plus consumption of media featuring trans/nonbinary people. While this measure contributed marginally to the maximal model ($\chi^2(1)=3.1$, $p=0.078$), the effect seems to come from the interaction of antecedent bias and specificity ($\chi^2(3)=8.3$, $p=0.039$). Second, *use* of singular *they* (affirmative, possible, negative responses) also contributed marginally to the maximal model (Fig 1), with participants reporting “yes” trending toward reading *themselves* faster than the other categories ($\chi^2(2)=5.7$, $p=0.059$).

Although no other significant results obtained from the social factors, there is suggestive evidence of what factors might detectably influence behaviour, given more refined methods. In addition, a small but significant effect of presentation order on reading time of *themselves* suggests short-term familiarity with *themselves* eases the associated processing costs (Fig 2) ($\chi^2(1)=12.8$, $p<0.001$). Future work will investigate whether this behaviour contributes to successful learning of singular personal *they*.

References:

- Conrod, K. (2017). "Dear Geoff". *Medium*. Retrieved on 7 December 2017 from: <https://medium.com/@kconrod/dear-geoff-bcd84286faa9>
- Doherty, A. and Conklin, K., 2017. How gender-expectancy affects the processing of "them". *The Quarterly Journal of Experimental Psychology*, 70(4), pp.718-735.
- Foertsch, J. and Gernsbacher, M.A., 1997. In search of gender neutrality: Is singular they a cognitively efficient substitute for generic he? *Psychological Science*, 8(2), pp.106-111.
- McLemore, K. A. (2015). Experiences with misgendering: Identity misclassification of transgender spectrum individuals. *Self and Identity*, 14(1), 51–74.

Figure 1:

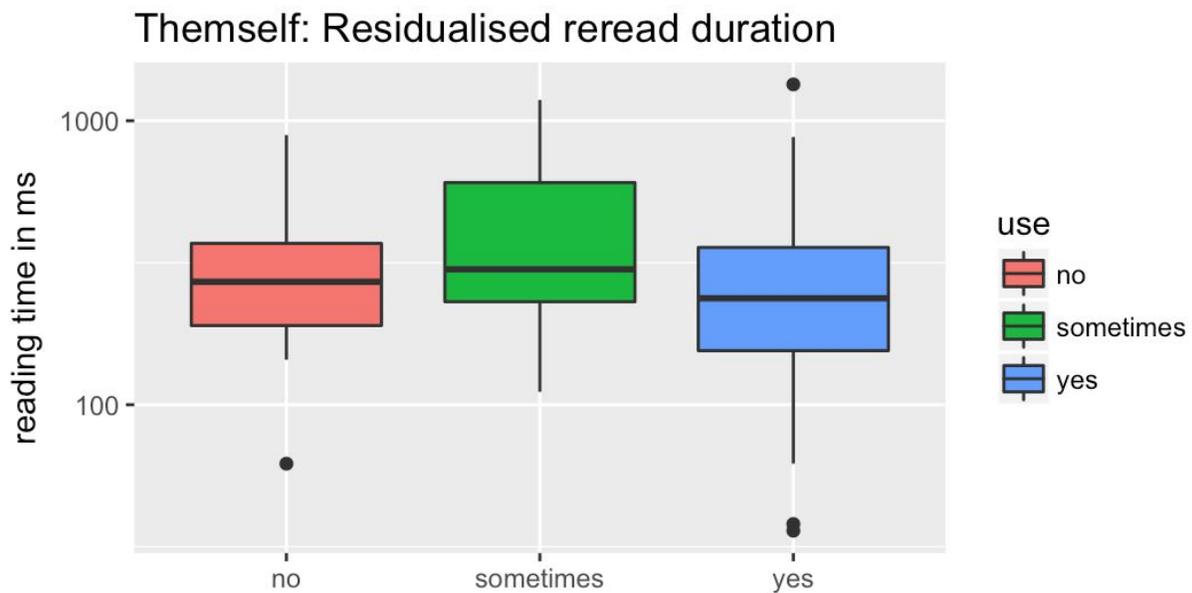
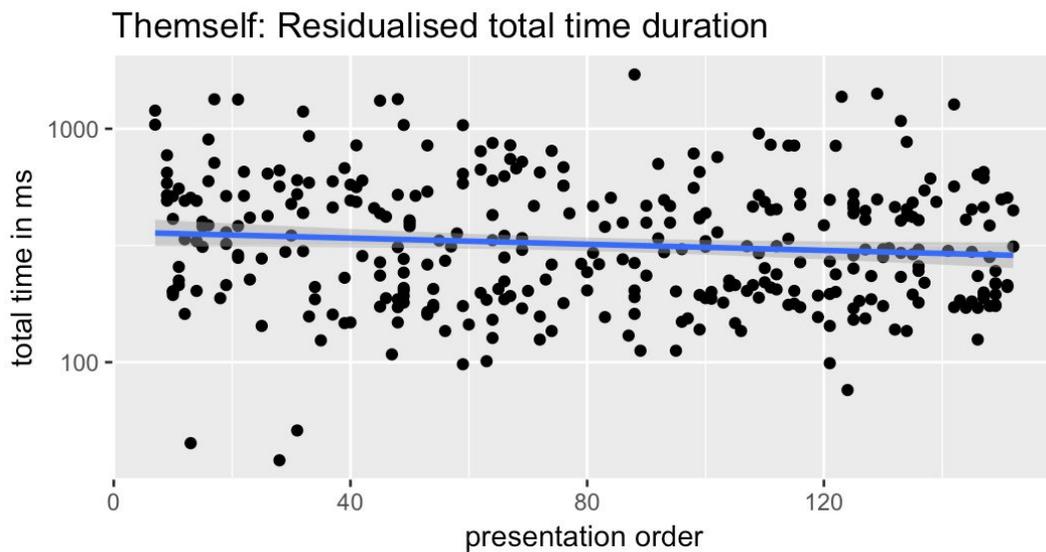


Figure 2:



Tailoring Referential Descriptions: Talking to Children, Adults, and Computers

Raheleh Saryazdi, Julie Bannon, & Craig G. Chambers (University of Toronto)

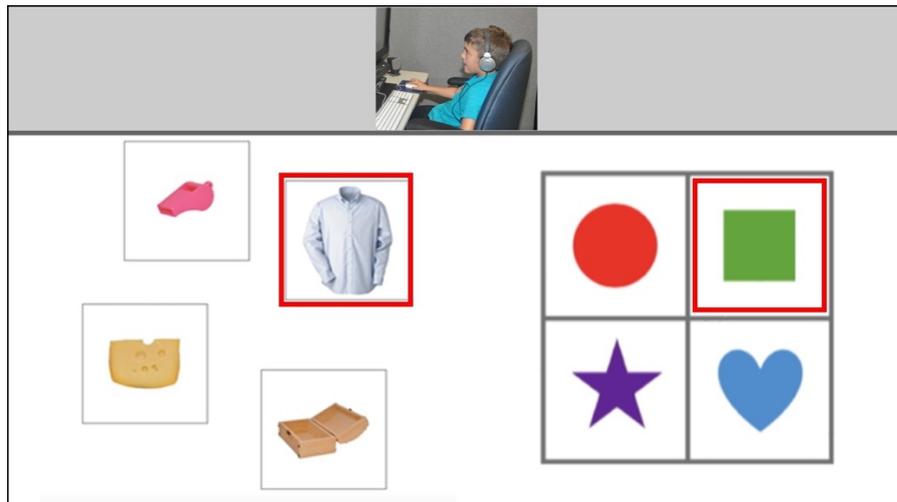
When designing referential expressions, speakers often make adjustments to adapt to addressees' information processing abilities. For example, previous research has shown speakers make various acoustic modulations when interacting with young children (infant-directed speech [1]), older adults (elderspeak [2]), and computers [3]. To what extent is the form and content of referring expressions also adjusted? Although numerous studies have examined how descriptions are tailored to the referential context (e.g., whether one or more objects from the same category are present), little work has conducted a concurrent investigation of the way that descriptions change according to audience type (although see [4]). The present study explores how audience type influences speakers' descriptive choices across referential contexts. More specifically, we compare how young adults adapt referential descriptions for a child or a computer compared to another young adult.

Twenty-four young adults were led to believe they were playing a communication game with three different addressees in an adjoining room: a young adult (20-year-old), a child (6-year-old), or a computer enabled with speech recognition software. Those who did not believe the cover story were replaced with new participants. At the beginning of each (randomized) block, players received confirmation from the experimenter that the other player was ready to play. The screen provided information regarding the player's age, and an image of the player (a clipart microphone indicated the speech recognition system). The image remained on top of the display for the entire block. On each trial, the display included a workspace with objects to be named on the left side and the intended location on the right side, which were sequentially cued. On critical trials, target objects were unique (Contrast Absent), or were accompanied by same-category competitors (Contrast Present).

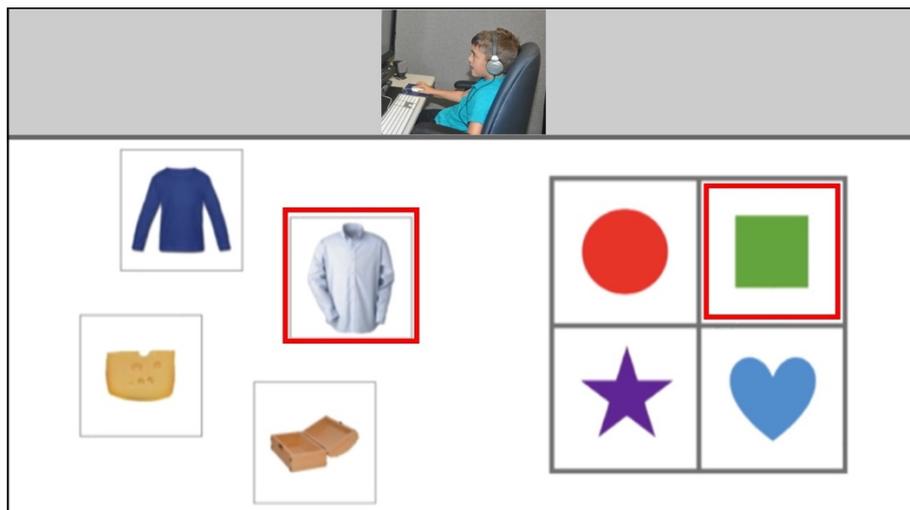
As expected, participants used more modifiers in the Contrast Present condition, where successful target identification required additional descriptors. Of particular interest, however, was whether naming patterns varied as a function of audience type. In terms of the amount of information provided, participants used more modifiers when speaking to a child than a young adult. Child-directed descriptions also reflected a higher incidence of colour modifiers relative to the other addressee groups. These adaptations may reflect inferences about vocabulary constraints in children and a strategy of drawing children's attention to salient features of target objects. Further, participants used significantly fewer modifiers with a computer addressee compared to young adults. Although previous work has reported differences for adults vs. computers [4], the pattern observed here suggests listeners either have unrealistic beliefs about the computer systems' ability to correctly identify an intended referent relative to humans, or that they are less motivated to ensure referential success. Intriguingly, the observed effects of audience type held across both Contrast Absent and Present conditions (no interaction), suggesting that the audience-type modulations were not tied to informativity in a straightforward way (i.e., in principle, no modification was needed in the Contrast Absent condition regardless of audience type). In addition, speech initiation times were consistent across audience types, indicating that the modulations in the computer and child conditions did not slow talkers' early-stage conceptualization processes.

The results show that, within some interesting limits, adult speakers attempt to spontaneously adapt referential strategies according to listener type, providing evidence for models of language interaction in which speakers consider addressees' processing abilities in addition to their informational needs.

Contrast Absent condition (showing cued target and destination),
with Child addressee



Contrast Present condition, with Child addressee



References:

- [1] Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20(1), 104-113, doi:10.1037/0012-1649.20.1.104
- [2] Kemper, S., Finter-Urczyk, A., Ferrell, P., Harden, T., & Billington, C. (1998). Using elderspeak with older adults. *Discourse Processes*, 25(1), 55-73, doi:10.1080/01638539809545020
- [3] Oviatt, S., MacEachern, M., & Levow, G. (1998). Predicting hyperarticulate speech during human-computer error resolution. *Speech Communication*, 24(2), 87-110, doi:10.1016/S0167-6393(98)00005-3
- [4] Schmader, C., & Horton, W. S. (2017). Conceptual effects of audience design in human-computer and human-human dialogue. *Discourse Processes*, doi:10.1080/0163853X.2017.1411716

INTEGRATION AND ANTICIPATION PROCESSES OF THE SPEAKER AND MEANING IN ADULTS WITH AND WITHOUT AUTISM SPECTRUM DISORDER: EVIDENCE FROM EYE-TRACKING AND ERPS

Mahsa Barzy (University of Kent), Jo Black (University of Kent), David Williams (University of Kent) & Heather J Ferguson (University of Kent)
mm951@kent.ac.uk

Evidence suggests that when comprehending language, people build mental models that include knowledge about the speaker. For instance, using event related potentials (ERPs), a semantic anomaly-type response (i.e. N400 effect) was observed when typically developing (TD) adults listened to sentences in which the content of the message mismatched with the voice of the speaker (van Berkum, et al., 2008). It is claimed that individuals with autism spectrum disorders (ASD) experience specific difficulties integrating information from the context to build pragmatic mental models while comprehending language (Happe, 1996). Therefore, we present two pre-registered experiments that examined whether adults with ASD exhibit comparable anticipation and integration processes for speaker and meaning as TD adults.

Experiment 1 employed the visual world paradigm, and tracked the timecourse of anticipatory biases about a speaker's meaning, based on their voice. Forty-eight participants (N=24 in each group, matched on gender, age and IQ) listened to sentences, in which the voice of speaker was either consistent or inconsistent with the intended message (e.g. "On my last birthday, I got an expensive electric shaver/car" in a child or an adult's voice), and concurrently viewed visual scenes that depicted these consistent and inconsistent objects alongside distractor objects. Participants were instructed to select the picture that best matched the audio description, and eye movements were recorded throughout. Behavioural results showed that for age and gender speaker types all participants were slower to select the correct object when it was inconsistent with the speaker's voice than when it was consistent. Eye-tracking results revealed a visual bias towards the object that was consistent with voice of the speaker group well before the disambiguating word onset. Hence, participants rapidly integrated the speaker's voice and used this to anticipate the content of forthcoming language. Growth curve analyses revealed a Group x cubic fit interaction, showing that the TD group fixated the target earlier than the ASD group (2240ms vs. 1800ms before disambiguation). The anticipatory bias in the TD group subsequently declined prior to a rapid increase (960ms before disambiguation), whereas the ASD group showed a consistent increase in target bias from 1800ms.

In Experiment 2 we recorded ERPs to explore how consistency between the speaker's voice and message influences integration processes. Forty-eight participants (N=24 in each group) listened to sentences of the same type as in Experiment 1 (e.g. "I tried to refresh my lipstick in front of the mirror" in a man or a woman's voice). EEG activity was recorded from 30 active electrodes, time-locked to the onset of the disambiguating target word, which was either consistent or inconsistent with the speaker's voice. A third sentence condition included a semantic anomaly (e.g. "I tried to refresh my seashell in front of the mirror"), and thus provided a baseline of anomaly detection effects on the N400 for comparison with speaker consistency effects. Results revealed an enhanced N400 for inconsistent sentences relative to consistent sentences, which was comparable to the N400 elicited by anomalous sentences. Further analyses revealed group differences in the topography of N400 effects, suggesting that different neural generators may be involved. Overall, these results show that contrary to previous suggestions of pragmatic dysfunction, people with ASD *are* sensitive to integration between speaker and meaning.