Gender differences in language use

Carl Vogel

Computational Linguistics Group
Trinity Centre for Computing and Language Studies
School of Computer Science and Statistics
Trinity College Dublin

Discourse, Technology and Translation

October 13, 2017

Digression: translation effects

- Compare large translation jobs with reference corpora using character and word n-gram similarity measures.
- Automatically identify the segments of the probe documents least similar to the reference corpus.
- Obtain a means of enabling large scale stylistic conformity checking for translation quality assurance, without a codified style.

Vogel, C., Lynch, G., Moreau, E., Sanchez, L. M., & Ritchie, P. (2013). Found in translation: computational discovery of translation effects. *Translation Spaces*, *2*(1), 81–104.

Gender differences in language use in dialogue

The Plan:

1. Gender differences in the use of English
2. Research questions
3. Method of analyzing the HCRC Map Task data
4. Results (yes, gender differences in English are evident)
5. Discussion

Conclusions

1. Recall that the original formulation of the Turing-test involved the question of whether a human would have the same accuracy in guessing the correct gender of a man and a woman where one is pretending to be the gender of the other as when one of the latter is replaced by a computer (Turing, 1950).
2. Understanding the role of gender in communication is important.
3. Gender differences in first person pronoun use are evident in the HCRC maptask corpus of dialogue.
4. Availability of eye contact matters.
5. Partner gender matters.
6. If the linguistic choices of authors and speakers reveals gender, what are the ramifications for translators and translation?

Evidence for gender differences in language use

- Newman, Groom, Handelman, and Pennebaker (2008) studied conversational data and fiction and found 37 linguistic categories for wich significant gender differences emerged
  19 were stable (differing in magnitude) across conditions, while 18 showed conditional reversed effects of gender (an *effect switch*).
  Pronouns use (more than females than by males) was stable.
- Argamon, Koppel, Fine, and Shimoni (2003) studied fiction and non-fiction and found females used first-person singular pronouns more than males in both fiction and non-fiction, and first-person plural pronouns more than males in non-fiction (but males using more first-person plural pronouns than females in fiction).
- Biber, Conrad, and Reppen (1998) argue that first and second person pronouns display involvement, that females display more involvement.
- Schmid (2015) studied the HCRC Map Task data, as a source that holds constant topic, looking at frequencies of a handful of items, but did not find significance for *I* or *you*.

Questions addressed here

1. Does dialogue role interact with gender and the use of *I* and *you* (cf. Schmid (2015))?
2. Does dialogue role interact with gender and the use of the larger category of pronouns (cf. Newman et al. (2008))?
3. Does dialogue role interact with gender and the use of first person singular and plural pronouns (cf. Argamon et al. (2003))?
4. Are there mitigating effects of the availability of eye-contact, familiarity of partners in a dyad, or partner gender, etc.?
5. Is there continuing support for an involvment account of differences (see Biber et al. (1998))?

Pronouns

Type	Instances
Demonstrative	*this*, *that*, *these*, *those*
Interrogative	*who*, *whose*, *which*, *whichever*, *whom*, *whomever*, *how*, *however*, *when*, *whenever*, *where*, *wherever*, *what*, *whatever*, *why*
First person singular	*i*, *me*, *myself*, *my*, *mine*
First person plural	*we*, *us*, *our*, *ours*, *ourselves*
Second person	*you*, *your*, *yours*, *yourself*, *yourselves*
Third Person	*they*, *them*, *their*, *theirs*, *it*, *he*, *she*, *him*, *her*, *his*, *hers*, *herself*, *himself*, *themselves*, *itself*, *its*, Demonstratives, Interrogatives

Data

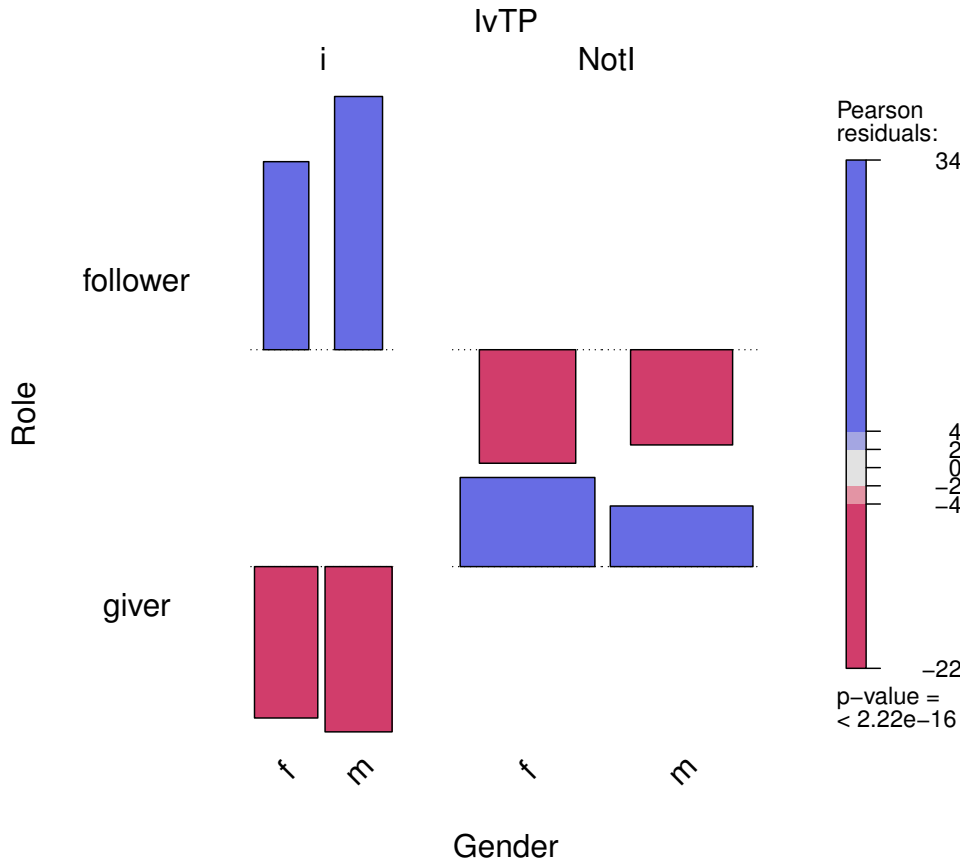
- The HCRC-Edinburgh-Glasgow map-task corpus provides transcriptions of two-person dialogues in which participants guide each other, using dialogue alone, in reproducing a path on maps that potentially are not identical.
  (For details, see: Anderson, Bader, Bard, Boyle, Doherty, Garrod, Isard, Kowtko, McAllister, Miller, Sotillo, Thompson, & Weinert, 1992)
- For any dialogue, one participant is an *information giver* (IG), communicating a path, and the other is an *information follower* (IF) reproducing the path on the partnered map.
- 64 native speakers of English (61 Scots) participated in four dialogues.
- 32 had the possibility of eye contact and the other 32 did not.
- Each participant played two games as IG and two as IF
- Each participant played in each role once with their familiar partner and once with an unfamiliar partner.

Data treatment

- 128 plain text files of the HCRC Map Task were processed by turn, recording the number of instances of pronouns and and total lexical tokens per line were counted.
- Counts per line were aligned with metadata of dialogue participants (ID, role, gender, age) experimental condition (availability of eye-contact, map used, ordinal position of the task, familiarity of partners, etc.)
- To demonstrate further effects we use association plots drawn from contingency tables, using the significance of Pearson standardized residuals to assess interactions that are significant (Friendly, 1999; Meyer, Zeileis, & Hornik, 2006).

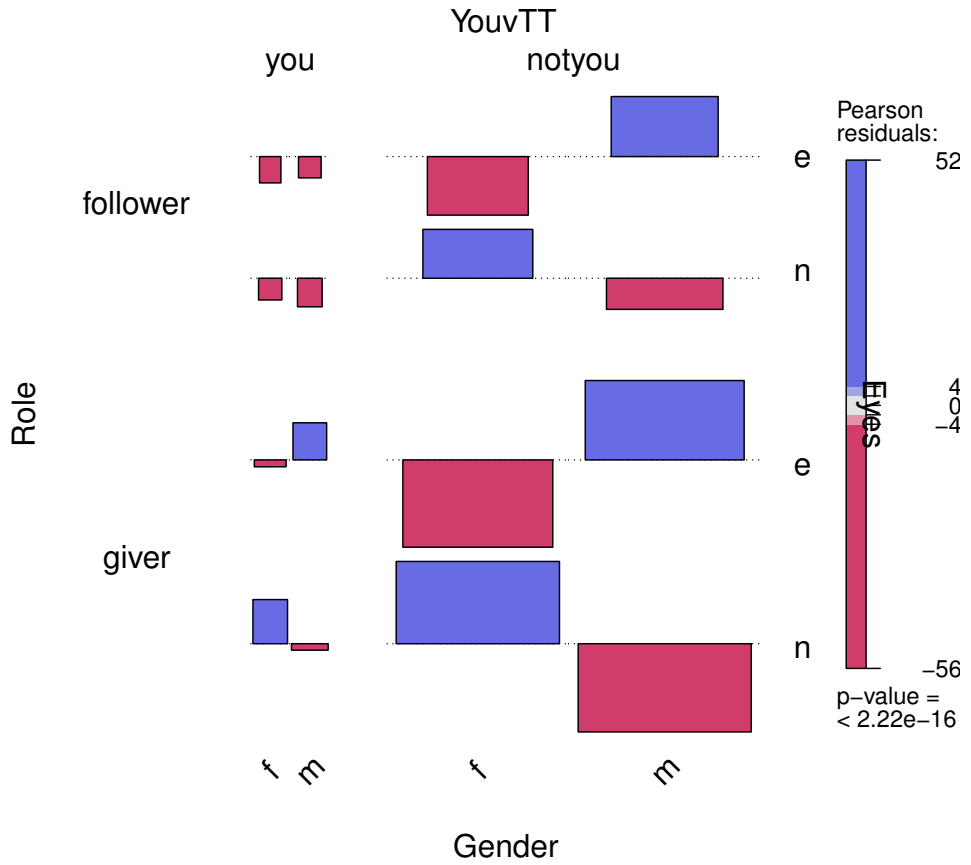
Interactions among *I* (vs. other pronouns), role and gender

- Information followers (IF) use *I* more than other pronouns than one would expect if there were no interaction between Role and *I* (Information givers, less).



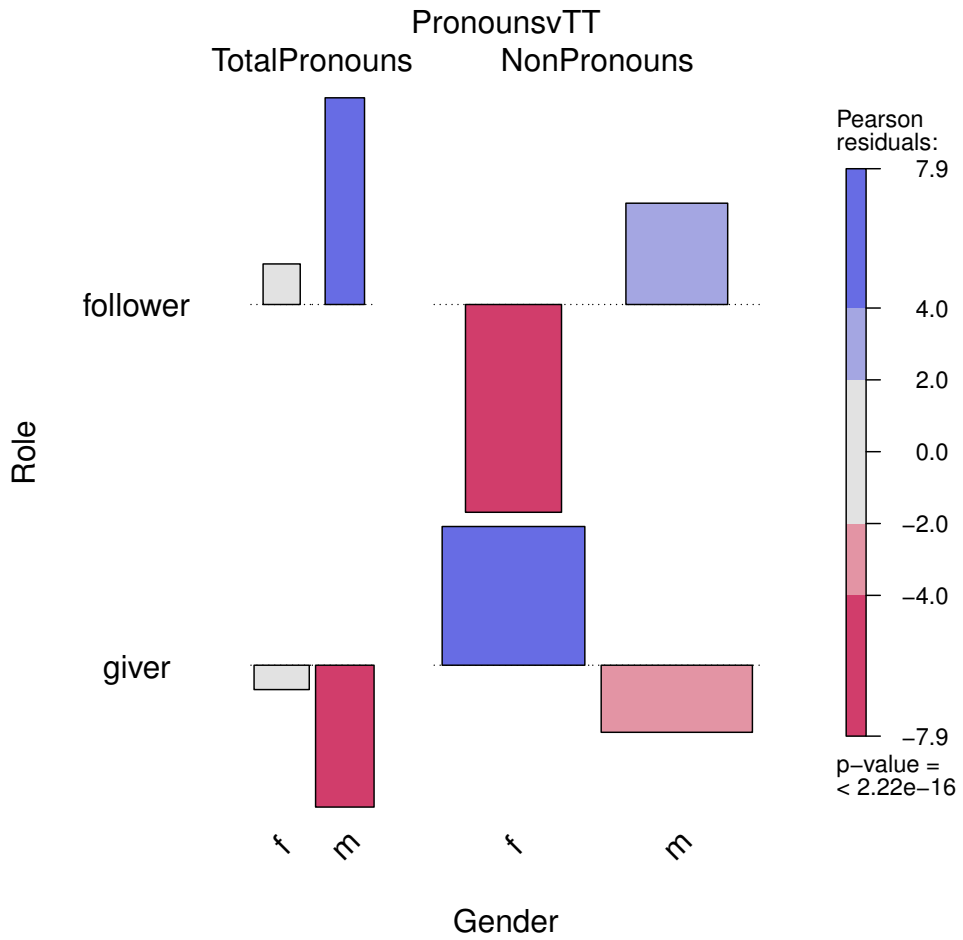
you (vs. other pronouns), role, gender and eye contact

- Information followers (IF) use *you* **less** than other pronouns than one would expect if there were no interaction between Role and *you* (Information givers, **more**), but there is a gender effect for IG that reverses with eye contact.



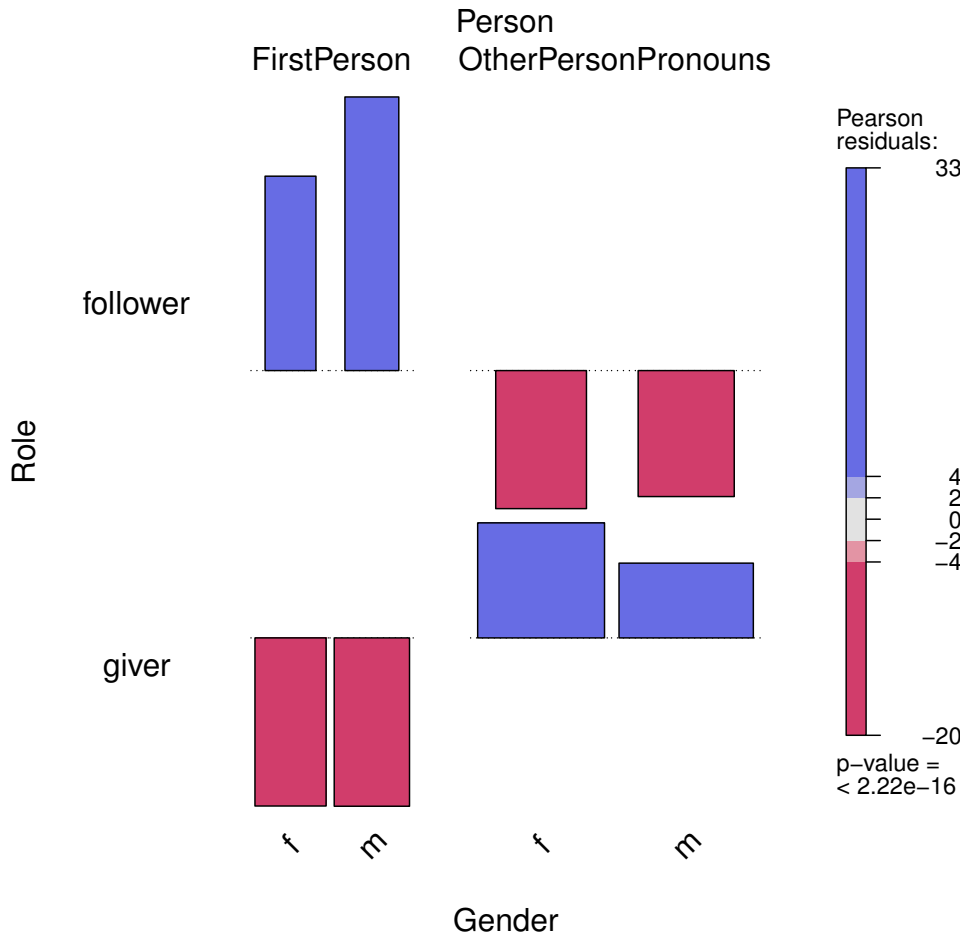
Pronouns vs. non-pronouns, role and gender

- Males, as IF, use pronouns **more** and as IG **less** than would be expected with no interation between pronouns and gender; observed values do not deviate significantly from expected values for women.



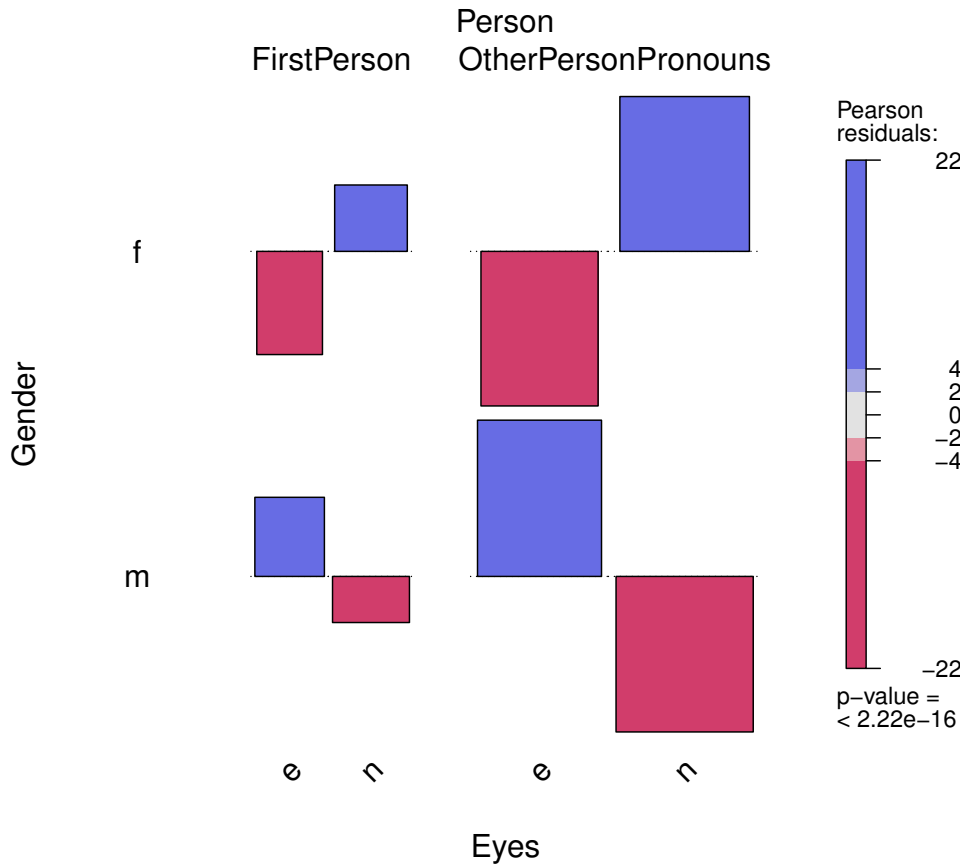
First person pronouns, role and gender

- Males and females both use first person pronouns **more** than would be expected and other pronouns **less** as IF; and in the reverse pattern in the IG role.



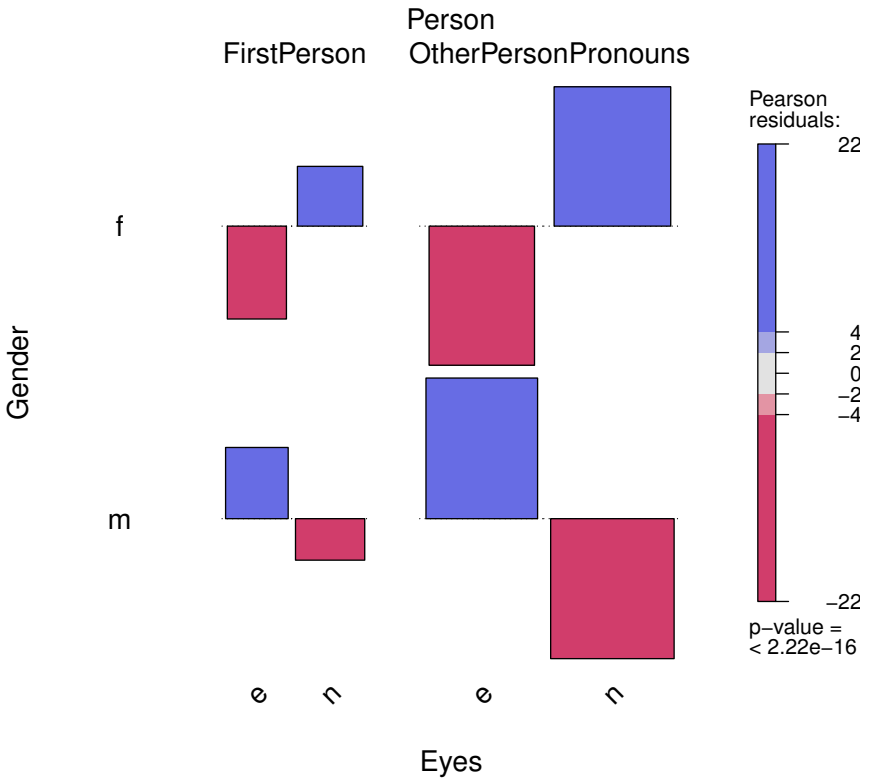
First person pronouns, gender and eye contact

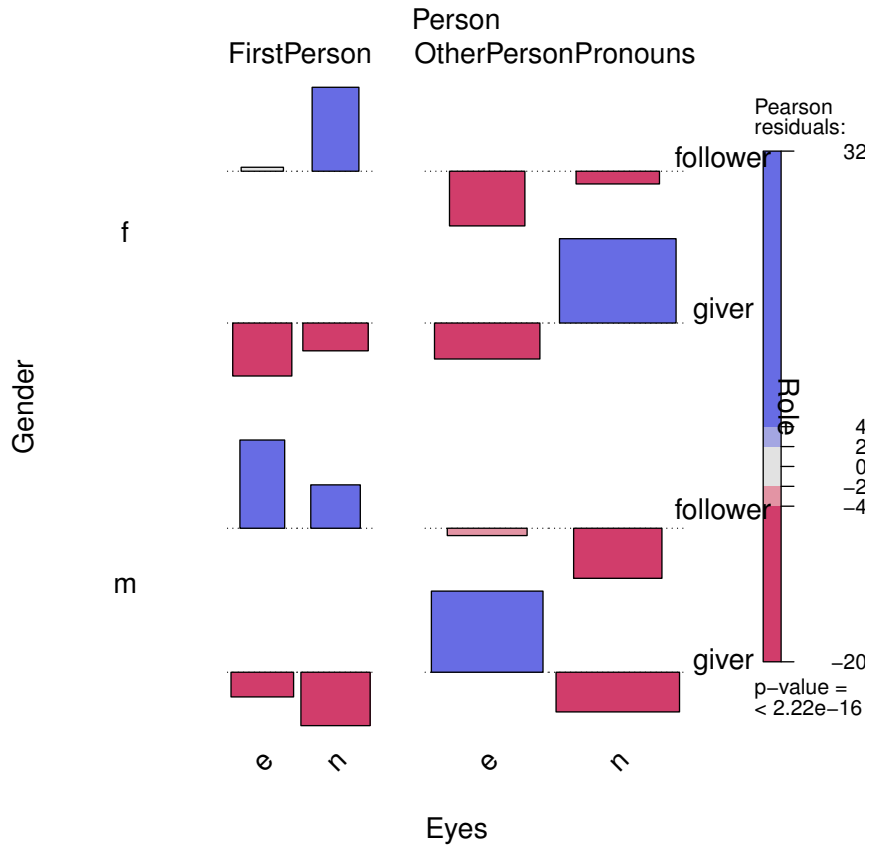
- Females use first person pronouns **more without eye contact** than would be expected and **less with eye contact**; males use first person pronouns **less without eye contact** and **more with eye contact** than would be expected without an interaction.



First person pronouns, gender, eye contact and role

- Females use first person pronouns **more without eye contact** and **less with eye contact** while males use first person pronouns **less without eye contact** and **more with eye contact** than would be expected.

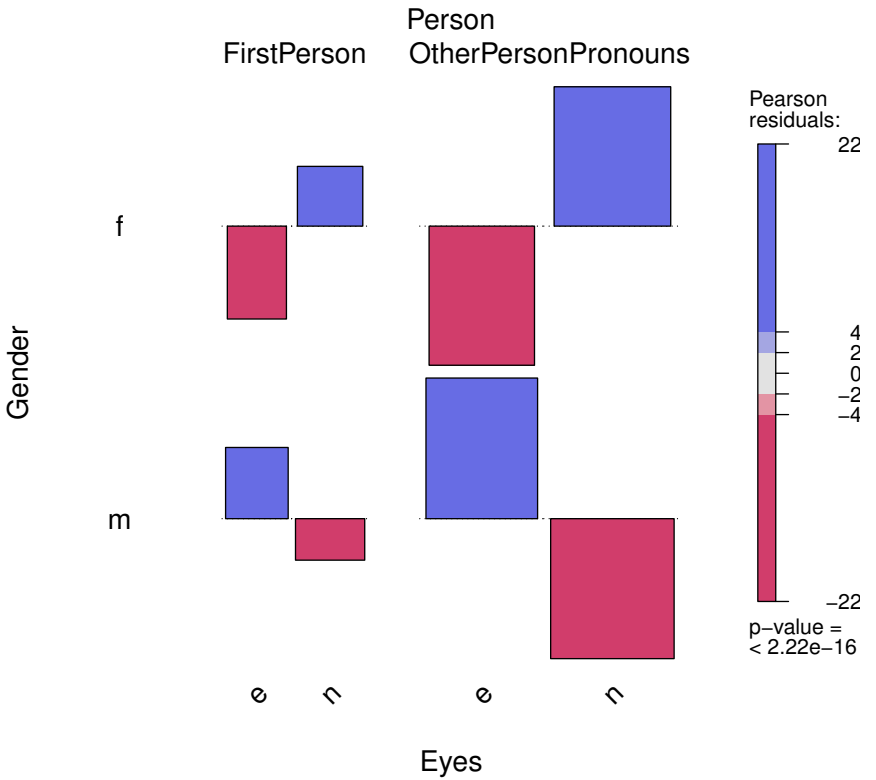


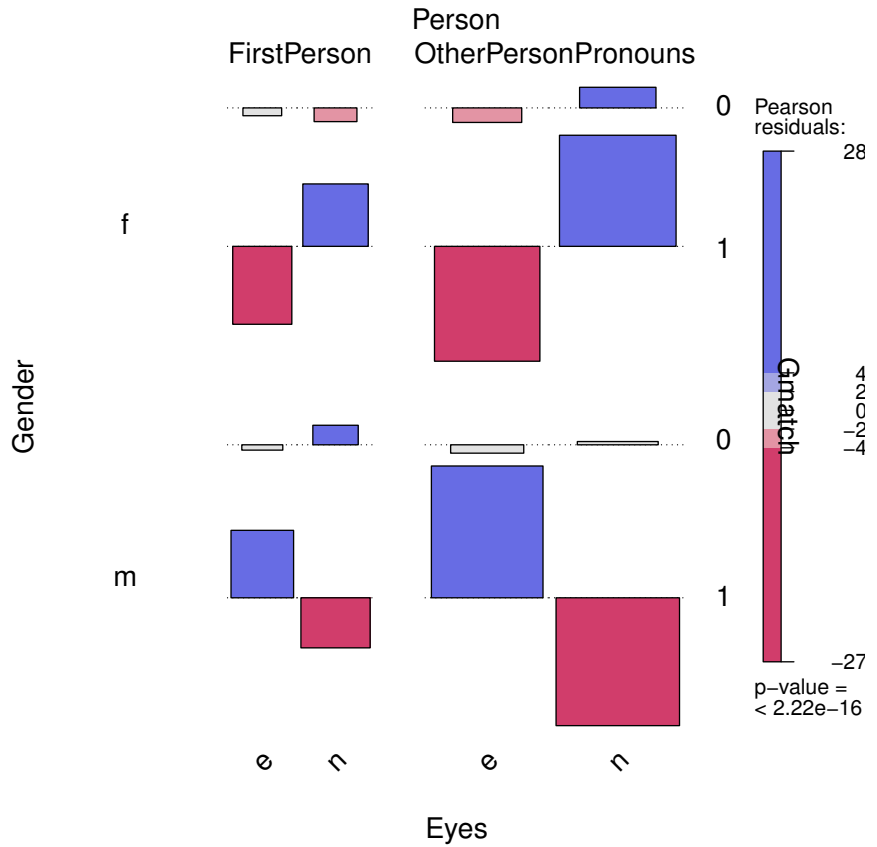


- **Dialogue role interacts.**

1st per. pronouns, gender, eye contact and partner gender

- Females use first person pronouns **more without eye contact** and **less with eye contact** while males use first person pronouns **less without eye contact** and **more with eye contact** than would be expected.

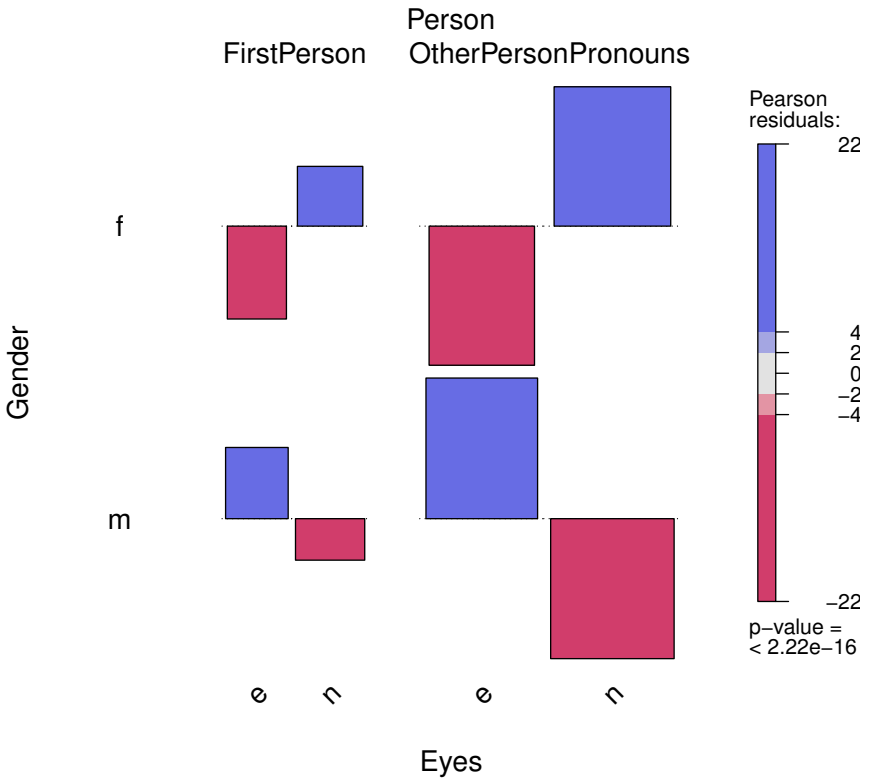


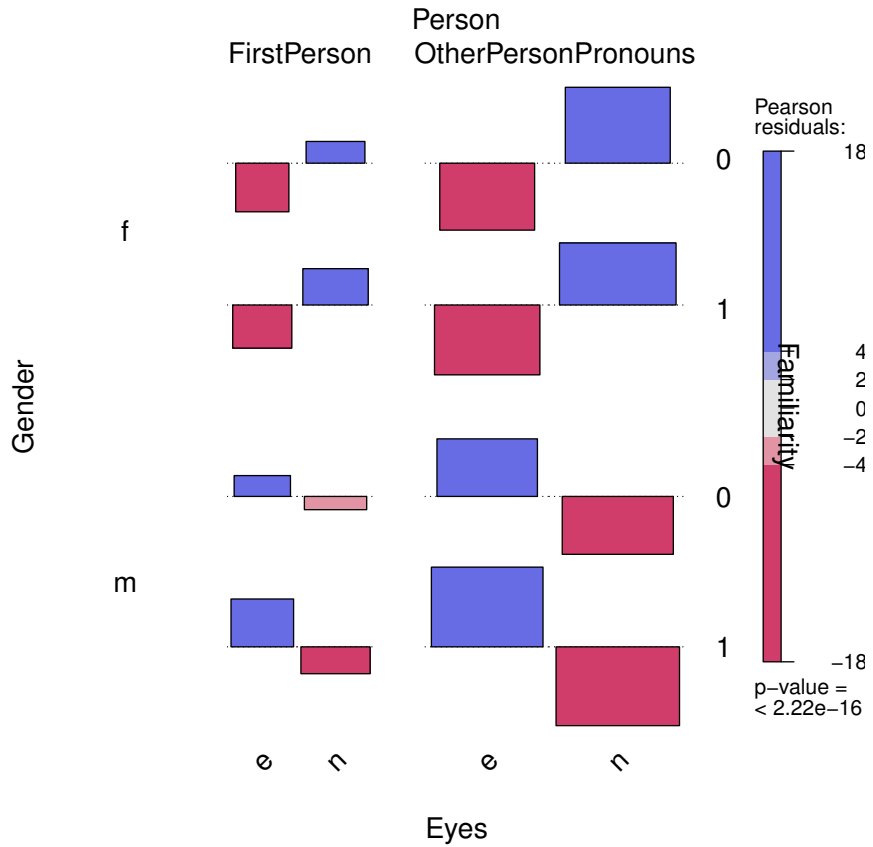


- **Partner gender mis-match diminishes the effect.**

1st per. pronouns, gender, eye contact and familiarity

- Females use first person pronouns **more without eye contact** and **less with eye contact** while males use first person pronouns **less without eye contact** and **more with eye contact** than would be expected.





- **Partner familiarity does not alter the effect.**

Discussion

1. Does dialogue role interact with gender and the use of *I* and *you*?
   For 'you', yes: there is an interaction with gender, role, and eye contact – as IG, females used you more without eye contact and less with eye contact than if there were no intereaction (males had the reverse pattern).
2. Does dialogue role interact with gender and the use of the larger category of pronouns?
   There is an effect of role that does not interact with gender: greater than expected pronoun use as IF and less as IG.
3. Does dialogue role interact with gender and the use of first person singular and plural pronouns?
   Yes – in interaction with other factors
4. Are there mitigating effects of the availability of eye-contact, familiarity of partners in a dyad or partner gender?
   Yes

Conclusions

1. Gender differences in first person pronoun use are evident in the HCRC maptask corpus.
2. Availability of eye contact matters
3. Partner gender matters
4. Recall that the original formulation of the Turing-test involved the question of whether a human would have the same accuracy in guessing the correct gender of a man and a woman where one is pretending to be the gender of the other as when one of the latter is replaced by a computer (Turing, 1950).
5. Understanding the role of gender in communication is important.
6. **If the linguistic choices of authors and speakers reveals gender, what are the ramifications for translators and translation?**

Finally

Thank you for your time and attention!!!

References I

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E. H., Doherty, G. M., Garrod, S. C., Isard, S. D., Kowtko, J. C., McAllister, J. M., Miller, J., Sotillo, C. F., Thompson, H. S., & Weinert, R. (1992). The hcrc map task corpus. *Language and Speech*, *34*(4), 351–366.

Argamon, S., Koppel, M., Fine, J., & Shimoni, A. R. (2003). Gender, genre, and writing style in formal written texts. *Text*, *23*(3), 321–346.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.

Friendly, M. (1999). Visualizing categorical data. In Sirken, M., Hermann, D., Schechter, S., Schwartz, N., Tanur, J. M., & Tourangeau, R. (Eds.), *Cognition and Survey Research*, chap. 20, pp. 319–348. John Wiley and Sons.

Meyer, D., Zeileis, A., & Hornik, K. (2006). The strucplot framework: visualizing multi-way contingency tables with vcd. *Journal of Statistical Software*, *17*(3), 1–48.

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: an analysis of 14,000 text samples. *Discourse Processes*, *45*(3), 211–236.

Schmid, H.-J. (2015). Does gender-related variation still have an effect, even when topic and (almost) everything else is controlled?. In Daems, J., Zenner, E., Heylen, K., Speelman, D., & Cuyckens, H. (Eds.), *Change of Pardigms–New Paradoxes: Recontextualizing Language and Linguistics*, Vol. 31 of *Applications of Cognitive Linguistics*, pp. 327–346. Berlin: De Gruyter Mouton.

Turing, A. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433–460.

Vogel, C., Lynch, G., Moreau, E., Sanchez, L. M., & Ritchie, P. (2013). Found in translation: computational discovery of translation effects. *Translation Spaces*, *2*(1), 81–104.