

Tracking the change in Hebrew possessive constructions using a blog corpus

Tal Linzen
New York University

Hebrew possessive constructions

Hebrew has two possessive constructions:

(1) Ordinary possession (OP):

hu shavar et ha-yad sheli
he broke ACCUSATIVE the-hand my

(2) Possessive dative (PD):

hu shavar li et ha-yad
he broke to.me ACCUSATIVE the-hand

The PD construction is a recent borrowing from European languages. While still considered somewhat substandard, it has been gradually gaining ground. In most European languages, the two constructions are not interchangeable: PD is used when the possessor is perceived to be **affected** by an event (König and Haspelmath 1998).

There has been some debate about the status of affectedness in the Hebrew PD (Landau 1999, Pykkänen 2002). Recent work has shown that while affectedness still plays a role in guiding speakers' choices between PD and OP, that role is definitely weaker than in the European prototype (Linzen, submitted).

Hypothesis: PD is undergoing historical change, from the European affectedness prototype to a general purpose possessive construction.

The possessor is most clearly affected when the possessed object is a **body part**. In addition, it is easy to determine in a corpus whether the possession is a body part or not — an quantifiable objective proxy for the subjective notion of affectedness.

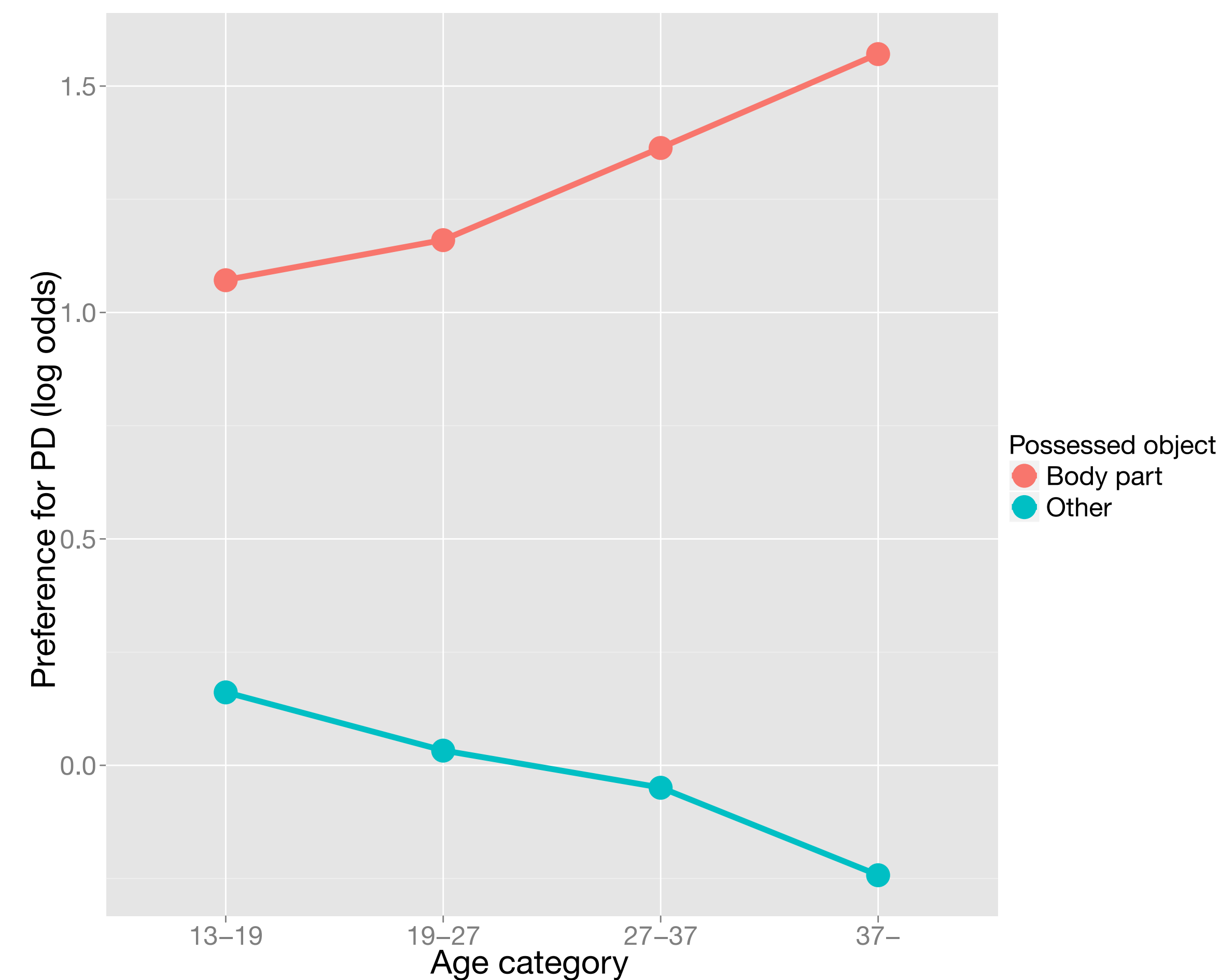


Figure 1: The difference between body part possessions and other possessions is more pronounced in older bloggers than in younger bloggers. If the trend continues at the same pace, in 30 years the distinction between the two classes of possessions will have no effect on the variation between the constructions.

Data set and analysis

- Possessive sentences were extracted from the Israblog corpus, which has **165 million tokens**
- The site encourages bloggers to specify their **age and gender**
- Heterogeneous mix of genres: important for studying substandard constructions
- There's a lot of polysemy in the dative construction, so **203 verbs** were manually selected for which the dative construction is unambiguously possessive
- The final dataset consisted of **49425 tokens of OP and PD** drawn from **1745 blogs**. The possessed object was automatically classified as either a body part or another type of object
- The choice between the two possessive constructions was analyzed using a **mixed effects logistic regression** in R with a random intercept and random slope for the body part predictor (structure justified by model comparison)

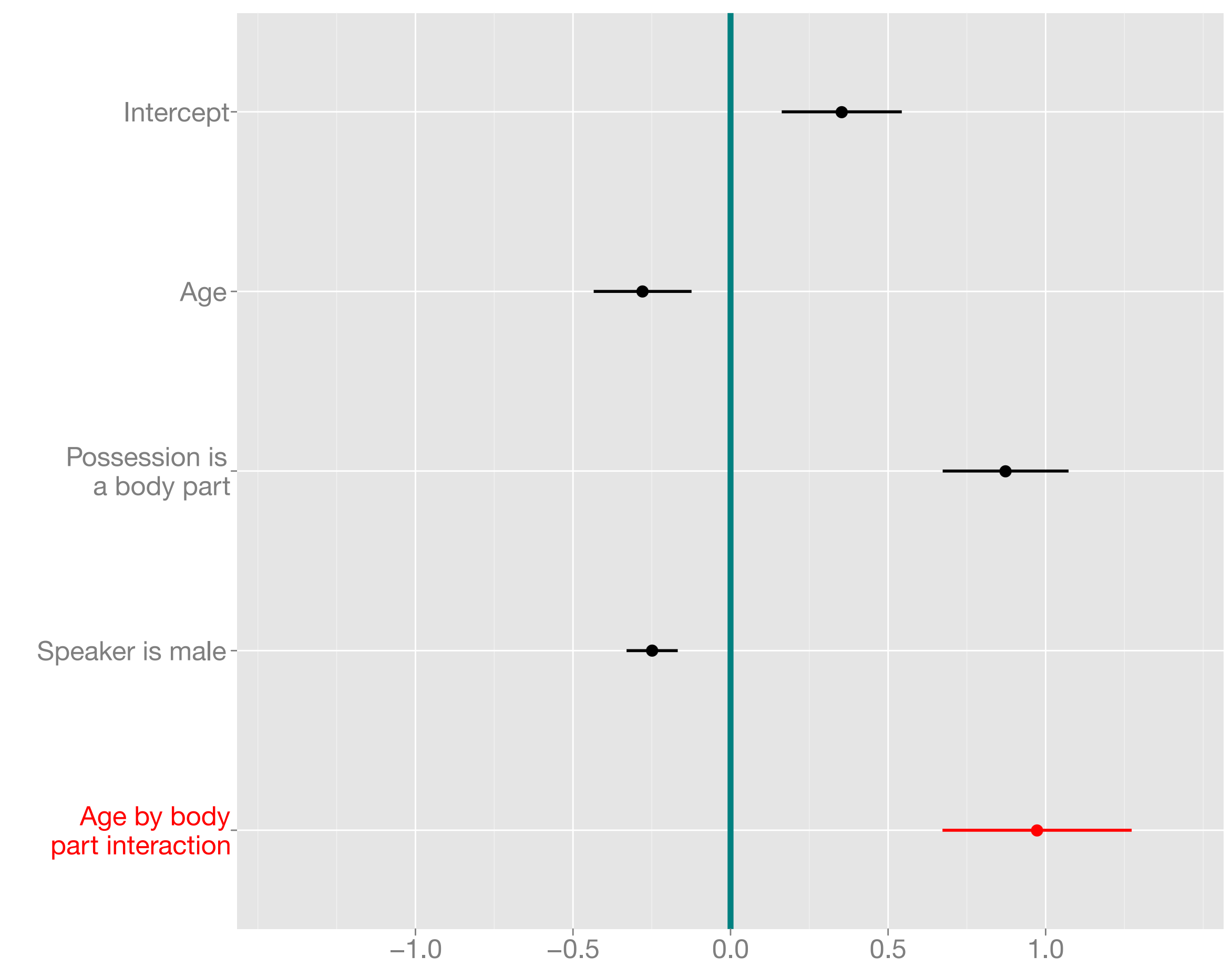


Figure 2: Results of the logistic mixed effect model (only significant effects). The critical interaction is shown in red: body parts have a stronger preference for PD in older speakers than in younger speakers.

Summary of results

- Younger bloggers are less sensitive than older bloggers to the association between body parts and PD
- Hebrew PD is drifting away from the European affectedness prototype and becoming a general possessive construction
- Age-marked blog platforms provide useful apparent-time corpora for studying language change

References

- Berman, Ruth. 1982. Dative marking of the affectee role: Data from Modern Hebrew. *Hebrew Annual Review* 6:35–59.
- Landau, Idan. 1999. Possessor raising and the structure of VP. *Lingua* 107:1–37.
- Linzen, Tal. submitted. Hebrew possessive datives: corpus evidence for the role of affectedness.
- König, Ekkehard, and Martin Haspelmath. 1998. Les constructions à possesseur externe dans les langues d'Europe. In *Actance et valence dans les langues de l'Europe*, ed. Jack Feuillet, 525–606. Berlin: Mouton de Gruyter.
- Pykkänen, Liina. 2002. *Introducing arguments*. Doctoral Dissertation, MIT.