Speakers extrapolate community-level knowledge from individual linguistic encounters
Anita Tobar Henríquez, Hugh Rabagliati, Holly Branigan, University of Edinburgh

Speakers vary their lexical choices depending on recent lexical processing, e.g., they tend to reuse the same words as their interlocutors (Brennan & Clark, 1996; Branigan et al., 2011). However, it is unclear how speakers' lexical choices are affected by community-level factors, e.g., whether the interlocutor is from their own speech community (*in-community partner*) or not (*out-community partner*). Indeed, we know very little about how speakers learn community-level linguistic knowledge. In three experiments, we examined (i) how speakers' referential choices varied depending on their partner's choices and speech community, and (ii) how speakers' extrapolation of their own choices to a subsequent partner was modulated by their partners' speech communities.

In Experiment 1, 160 Spanish participants completed two sessions of a picture-naming task, where they took turns with a confederate to select and name a target. They encountered different confederates in each session. Experimental items comprised targets with both a high-frequency and a low-frequency label in participants' linguistic community (e.g., *patata* [potato] vs *papa* [spud]). In Session 1, the confederate named targets before the participant, using only low-frequency labels, and we measured participants' tendency to reuse such labels (**Lexical Entrainment**). In Session 2, only participants named the targets, and we measured participants' tendency to reuse the entrained terms they had used in Session 1 (**Maintenance of Entrained Terms**). As shown in Figure 1, we manipulated participants' beliefs about their confederates' linguistic community using a 2x2 design: In Session 1, the confederate was either an *in-community partner* from Spain or an *out-community partner* from Latin America (i.e., **First Partner's Community**); in Session 2, the confederate was either from the same community as the first partner or not (**Second Partner's Community**). Experiment 2 reproduced Experiment 1 in Mexican population (N=160). In Session 1, the confederate was either an *in-community partner* (Mexico) or an *out-community partner* (Argentina), and we measured **Lexical Entrainment**; in Session 2, the confederate was either from the same community as the first partner or not, and we measured **Maintenance of Entrained Terms**. In addition, Experiment 3 tested the effects of perceived linguistic status on entrainment and maintenance in 80 Mexican participants. In Session 1, the confederate was either a high-status out-community partner (Spain) or a low-status out-community partner (Argentina), and we measured **Lexical Entrainment**; in Session 2, all participants interacted with a middle-status in-community (Mexican) partner, and we measured **Maintenance of Entrained Terms**.

In Experiment 1 (Figure 2), disfavoured terms were used significantly more in Session 1 (50%[30%]) than in a spontaneous naming task (4%[6%]; V=0, p<.0001), suggesting a **Lexical Entrainment Effect**. But lexical entrainment was not affected by **First Partner Community** (β =.038, SE=.15, z=.25, p>.05): Participants entrained to similar rates with a partner from another speech community (52%[29%]) or from their own speech community (49%[32%]). In Session 2, however, participants generalised their expressions from Session 1 based on their confederates' communities. There was a significant interaction between **First Partner's Community** and **Second Partner's Community** (β=-.3, SE=.13, z=-2.3, p=.02): Participants who first entrained to an out-community partner maintained those entrained terms less often with an in-community partner in Session 2 (57% [SD=32%]) than with an out-community partner (71%[21%]; β=-.38, SE=.18, z=-2.2, p=.027); in contrast, participants who entrained to an in-community partner maintained terms to similar extents with an in-community partner as an out-community partner (72%[30%] vs. 79%[20%]; β=0.21, SE=.19, z=1.1, p>.05). Experiment 2 (Figure 3) replicated this pattern of results in Mexican participants and Experiment 3 (Figure 4) showed that linguistic status had no effect on either lexical entrainment (β=.09, SE=.24, z=.38, p>.05) or maintenance (β=.24, SE=.23, z=1.02, p>.05), suggesting that our results were driven by differences in confederates' communities, rather than linguistic status.

These results suggest that speakers encode speech community information during language processing and store that information to inform future contexts of language use, even when such community information has not affected speakers' language use during that particular linguistic encounter. Critically, they show that speakers learn community-level knowledge by extrapolating linguistic information from individual-level experiences.
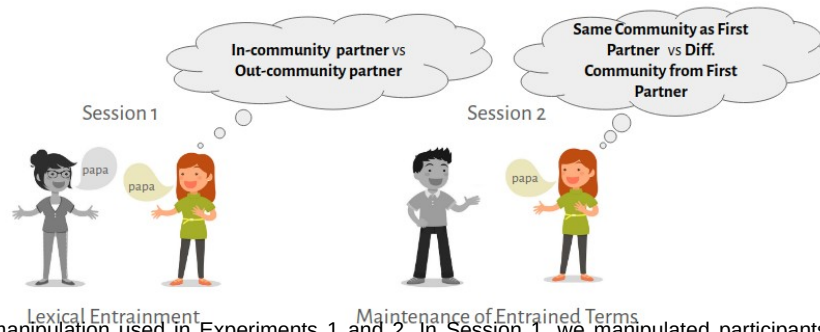
Figure 1. Experimental manipulation used in Experiments 1 and 2. In Session 1, we manipulated participants' beliefs about whether the confederate was either an in-community partner or an out-community partner (First Partner's Community: In-community partner vs out-community partner). In Session 2, we manipulated participants' beliefs about whether the confederate was either from the same community as the first partner or not (Second Partner's Community: Same Community as First Partner vs Different Community from First Partner).
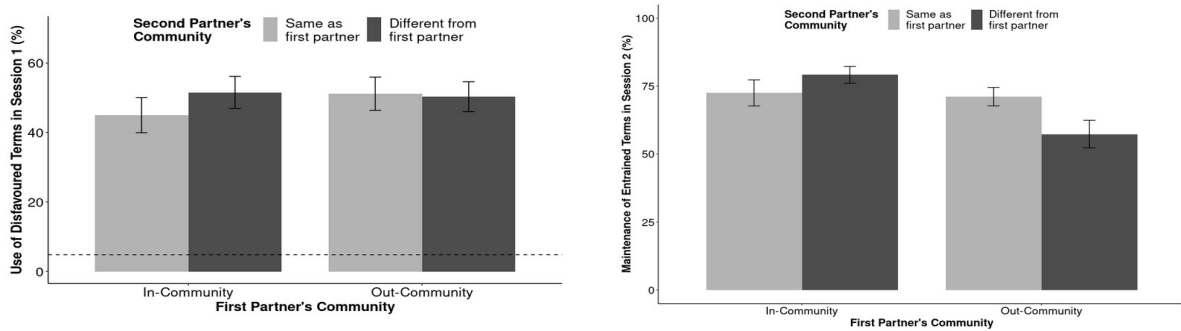

Figure 2. Experiment 1. **Left:** Mean and standard error of the percentage of use of disfavoured terms in **Session 1** (y-axis) by First Partner's Community (x-axis) and Second Partner's Community (colour-coded). The horizontal dashed line represents the baseline, i.e., the mean of percentage of use of disfavoured terms in a spontaneous naming task. **Right:** Mean and standard error of percentage of maintenance of disfavoured terms used in Session 1 during **Session 2** (y-axis), by First Partner's Community (x-axis) and Second Partner's Community (colour-coded).
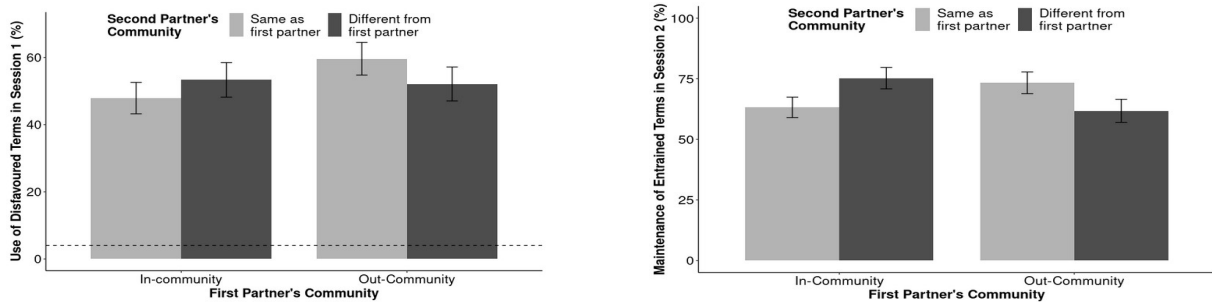

Figure 3. Experiment 2. **Left:** Mean and standard error of the percentage of use of disfavoured terms in **Session 1** (y-axis) by First Partner's Community (x-axis) and Second Partner's Community (colour-coded). The horizontal dashed line represents the baseline, i.e., the mean of percentage of use of disfavoured terms on a spontaneous naming task. **Right:** Mean and standard error of percentage of maintenance of disfavoured terms used in Session 1 during **Session 2** (y-axis), by First Partner's Community (x-axis) and Second Partner's Community (colour-coded).

Figure 4. Experiment 3. **Left:** Mean and standard error of the percentage of use of disfavoured terms in **Session 1** (y-axis) across First Partner's Community (x-axis). The dashed line represents the mean of percentage of use of disfavoured terms on the pretest. **Right:** Mean and standard error of percentage of maintenance of disfavoured terms used in Session 1 during **Session 2** (y-axis) across First Partner's Community (x-axis).

References

Branigan, H. P., Pickering, M. J., Pearson, J., & McLean, J. F. (2010). Linguistic alignment between people and computers. Journal of Pragmatics, 42(9), 2355-2368.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(6), 1482-1493.