Supplementary Information for ‘Acquiring the lexicon and grammar of universal kinship’

Joe Blythe, Jeremiah Tunmuck, Alice Mitchell, and Péter Rácz

1. NOMINAL EXPERIMENT

1.1. STRUCTURE

Each participant responds to a tailor-made list of stimuli.

The aim of the experiment was to show kinship relations to each participant in such a way that we can test the effect of an increase in genealogical distance, an increase in the complexity of kinship mergers, along with the effect of task type (apposite comprehension questions / inapposite comprehension questions / production / altercentric production). We also wanted to sample participants from four age groups, keeping the gender ratio across groups relatively balanced.

Since each such kinship relation was captured by a real-life referent, and since the population of our fieldwork site, Wadeye, is relatively small, it was not possible to completely balance the number of question types per participant, or the number of participants per age group.

Below we see the stimuli list for the first participant in our data.

Trial n	Task name	Term	Question type	Genealogical distance	Kinship distance
1	comprehension	B	polar: apposite	1	1
2	comprehension	FB > F	polar: apposite	2	1
3	comprehension	M	polar: apposite	1	1
4	comprehension	FBD > FD > Z	polar: apposite	3	1
5	comprehension	FBD > FD > Z	polar: inapposite	3	1
6	comprehension	FFB > FF	polar: apposite	3	2
7	comprehension	FM	polar: inapposite	2	2
8	comprehension	FMZSD > FMZSD > FBD > FD > Z	polar: inapposite	5	1
9	comprehension	FMZSD > FMZSD > FZS	polar: inapposite	5	3
10	comprehension	FMBS	polar: apposite	4	4
11	comprehension	FZ	polar: inapposite	2	2
12	comprehension	MB	polar: inapposite	2	2
13	comprehension	FMZ > FM	polar: apposite	3	2
14	comprehension	MF	polar: apposite	2	2
15	comprehension	MMBD	polar: inapposite	4	4
16	comprehension	MMBD	polar: apposite	4	4
17	comprehension	MMBDS	polar: apposite	5	5

18	comprehension	FZH > MB	polar: apposite	3	2
19	production	FMBDS	content	5	5
20	production	FMZS > FMS > FB > F	content	4	1
21	production	MMBS	content	4	4
22	production	FBW > FW > M	content	3	1
23	production	FZD	content	3	3
24	production	FMZD > FMD > FZ	content	4	2
25	altercentricity	B	content	1	1
26	altercentricity	FZ	content	2	2
27	altercentricity	FMZSW > FMSW > FBW > FW > M	content	5	1
28	altercentricity	FMZ > FM	content	3	2
29	altercentricity	FMZS > FMS > FB > F	content	4	1

The tasks start easy enough ('is this your brother?') and get progressively harder, ending with 'what would this person call you?' [with a picture of the participant's father's mother's sister]. In the comprehension task, initial questions are more likely to be apposite, later questions are more likely to be inapposite.

The stimuli for each participant follow this general structure, with slight variation in the number of question types. We illustrate this with the across-participant variation in (i) task type and (ii) merger type below.

Participant ID	comprehension	production	altercentricity
Participant 1	18	6	5
Participant 2	18	6	3
Participant 3	17	6	3
Participant 4	17	6	3
Participant 5	18	5	3
Participant 6	17	6	4
Participant 7	17	6	4
Participant 8	17	6	4
Participant 9	17	6	4
Participant 10	18	6	5
Participant 11	18	6	5
Participant 12	18	6	5
Participant 13	18	6	5
Participant 14	17	6	3
Participant 15	18	6	5
Participant 16	18	6	5
Participant 17	17	6	5
Participant 18	20	6	6

Participant 19	20	6	6
Participant 20	20	6	6
Participant 21	20	6	6
Participant 22	20	6	6
Participant 23	14	6	6
Participant 24	20	6	6

Participant ID	0 same-sex sib merger(s)	1 same-sex sib merger(s)	2 same-sex sib merger(s)
Participant 1	16	9	4
Participant 2	17	6	4
Participant 3	16	6	4
Participant 4	16	6	4
Participant 5	16	7	3
Participant 6	11	14	2
Participant 7	11	14	2
Participant 8	11	14	2
Participant 9	11	14	2
Participant 10	15	9	5
Participant 11	15	9	5
Participant 12	15	9	5
Participant 13	15	9	5
Participant 14	13	9	4
Participant 15	16	9	4
Participant 16	15	10	4
Participant 17	14	10	4
Participant 18	18	10	4
Participant 19	18	10	4
Participant 20	18	10	4
Participant 21	18	10	4
Participant 22	18	10	4
Participant 23	13	10	3
Participant 24	17	11	4

The minimum number of comprehension questions for a participant is 14, the maximum is 20. The minimum number of referents with a relationship including two same-sex sibling mergers to the participant is 2, the maximum is 5.

Other similar asymmetries subsist in the stimuli, and, as a consequence, in the data. Same-sex sibling mergers are particularly illustrative: some children have a big enough family so that a larger number of these referents can be located, while some do not. The reader can use the data set for similar cross-references.

1.2. METHODS

The number of question types, then, varies slightly across participants. Referents are ‘recycled’: For example, Referent 11 is a father’s sister to some, a mother’s sister’s daughter to others, and a father’s mother’s sister’s daughter to yet others, and so on.

This is where hierarchical models shine: we are able to assume that participants and referents will have varying success rates not captured by our predictors – we can estimate random intercepts. Ideally, a participant might be more or less successful in answering specific types of questions – which we could estimate using random slopes – and this rate of success might be more or less variable depending on e.g. the amount of questions of a given type (a participant answering 20 comprehension questions has more room to vary than one answering 14 such questions). However, data structure and the amount of data available do not favour more complex model fits.

Many of our predictors are ordered categories. Task type in the nominal experiment is a good example: we expect that comprehension is easier than production and that production is easier than altercentric production. When estimating the effects of ordered predictors, polynomial relationships between predictor and outcome are also considered. Given data scarcity, we only report and discuss linear (L) effects in the paper.

Models were fit using brms and stan. We used weakly informative priors for beta and the intercept. In these cases, the weak priors are more conservative than non-informative priors, but results do not change substantially based on prior choice. Following advice by Gelman et al 2008, we specified $\text{student_t}(1,0,2.5)$ as our weak prior. We followed the WAMBS checklist in model fitting diagnostics (following van de Schoot 2019).

Below, we list the models and figures reported in the paper for the *nominal experiment*. The *code* for these models and figures is visible in the Rmd version of this Supplementary Information.

We use line segments to visualise model predictions and bar charts to plot aggregated raw data. We did not use violin plots because, in many cases, the small number of trials per participant result in a small number of unique averages, making violin plots a strain on the eye. Instead, we show the total counts of correct and incorrect answers per question types. The height difference between the bars captures participant accuracy for a given question type, while the difference between these differences captures how accuracy varies across question types. We also provide means and binomial standard errors on separate helper charts above the main charts (it makes sense when you look at it).

1.3. MODEL 1

Correct comp ~ Age band + Apposite + Kin distance + (1 | Identifier) + (1 | Referent_ID)

	Estimate	Est.Error	Q2.5	Q97.5
Intercept	1.88	0.55	0.89	3.06
Age.Band.L	1.37	0.56	0.28	2.51
Age.Band.Q	-0.58	0.57	-1.78	0.49
Age.Band.C	0.22	0.55	-0.86	1.33
AppositeTRUE	1.05	0.52	0.05	2.09
Kin.distance.L	-1.32	0.71	-2.70	0.06
Kin.distance.Q	1.46	0.78	0.05	3.14
Kin.distance.C	0.60	0.55	-0.48	1.73
Kin.distanceE4	0.17	0.57	-0.92	1.27

Figure 1

Figure 2 **1.4. MODEL 2**

Correct ~ Age band + Task type + Genealogical distance + Same sex sibling merger + Step kin merger + Half-sibling merger + (1 | Identifier) + (1 | Referent ID)

	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.83	0.50	-0.12	1.85
Age.Band.L	1.37	0.59	0.23	2.54
Age.Band.Q	-0.13	0.57	-1.28	1.00
Age.Band.C	0.55	0.56	-0.58	1.63
Task2.L	-3.25	0.48	-4.26	-2.36
Task2.Q	-0.17	0.37	-0.92	0.55
Genealogical.distance.L	-2.21	0.72	-3.64	-0.84
Genealogical.distance.Q	0.27	0.53	-0.70	1.36
Genealogical.distance.C	-0.04	0.46	-0.98	0.85
Genealogical.distanceE4	0.23	0.46	-0.68	1.15
Same.sex.sibling.merger.L	1.77	0.75	0.34	3.33
Same.sex.sibling.merger.Q	-0.25	0.41	-1.07	0.55
Step.kin.merger1	-0.72	0.60	-1.93	0.42
Half.sibling.merger.L	-0.64	0.89	-2.40	1.06
Half.sibling.merger.Q	0.48	0.48	-0.45	1.43

Figure 3 **Figure 4**

Figure 5 **Figure 6**

2. KINTAX EXPERIMENT

2.1. STRUCTURE

The kintax experiment depicts abstract situations using simple line drawings and an accompanying sound clip describing the depicted activity. The participants have to choose the appropriate pair of referents for each situation. Since Murrinhpatha kintax marks whether referents are related to *each other* as opposed to the speaker / listener, there was no need to create individual stimulus lists. As a result, the stimuli are balanced across participants, with 3 gender, 3 number, 7 sibling/non-sibling and 7 cross-/parallel-cousin trials per participant.

2.2. METHODS

We grouped the data under participants (estimating a participant random intercept). This is also a small data set so similar caveats apply as in the nominal experiment. Exploratory data analysis did not indicate that further grouping factors were relevant.

Model 3

Correct ~ Age band + Condition + (1 | Identifier)

	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.97	0.23	0.52	1.44
Age.bands.L	1.17	0.26	0.68	1.68
Age.bands.Q	-0.01	0.26	-0.53	0.50
Age.bands.C	0.53	0.26	0.02	1.04
Age.bandsE4	-0.07	0.26	-0.58	0.42
Conditiondualvpaucal	0.10	0.30	-0.48	0.69
ConditioncrossMvparallelMcousin	-0.47	0.25	-0.96	0.02
ConditionsiblingvnonMsibling	-0.11	0.25	-0.63	0.39

Figure 7

Figure 8

Figure 9

2.3. MODEL 4A

Correct ~ Age band + Condition + (1 | Identifier)

2.4. MODEL 4A

Correct ~ Average accuracy in nominal experiment + Condition + (1 | Identifier)

	Estimate	Est.Error	Q2.5	Q97.5
Intercept	0.49	0.34	-0.17	1.15
Age.bands.L	0.86	0.44	-0.05	1.74
Age.bands.Q	-0.44	0.42	-1.31	0.40
Age.bands.C	0.23	0.41	-0.61	1.02
Conditiondualvpaucal	0.34	0.40	-0.46	1.11
ConditioncrossMvparallelMcousin	-0.44	0.34	-1.11	0.20
ConditionsiblingvnonMsibling	0.16	0.34	-0.51	0.82
	Estimate	Est.Error	Q2.5	Q97.5
Intercept	-1.15	0.79	-2.76	0.37
kinap.av	2.43	1.12	0.27	4.71
Conditiondualvpaucal	0.33	0.39	-0.43	1.10
ConditioncrossMvparallelMcousin	-0.43	0.33	-1.07	0.23
ConditionsiblingvnonMsibling	0.16	0.33	-0.47	0.79

References

- DEPAOLI, SARAH.; and RENS VAN DE SCHOOT. 2017. Improving transparency and replication in Bayesian statistics: The WAMBS-Checklist. *Psychological methods* 22.240–261. doi: [10.1037/met0000065](https://doi.org/10.1037/met0000065).
- GELMAN, ANDREW.; ALEKS JAKULIN.; MARIA GRAZIA PITTAU.; and YU-SUNG SU. 2008. A weakly informative default prior distribution for logistic and other regression models. *The annals of applied statistics* 2.1360–1383. doi: [10.1214/08-AOAS191](https://doi.org/10.1214/08-AOAS191).

Data

Anonymised raw data and code have been placed in the public repository <https://zenodo.org/record/3732882>.