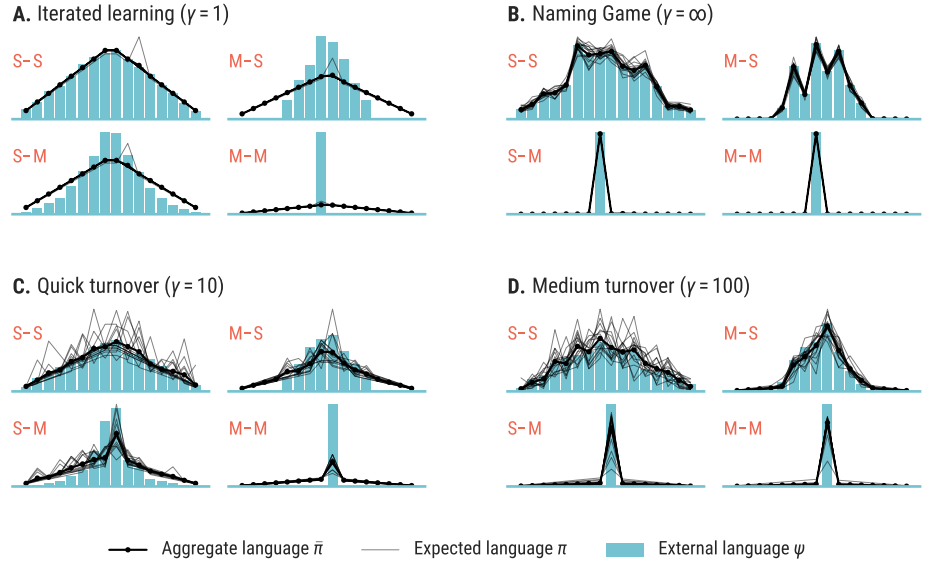# Conlusions

Bas Cornelissen

August 9, 2017

What can models learn us about language evolution? That is the question this thesis set out to address. Without any introduction, it suggests a fairly philosophical thesis, but we found nothing of the sort — deliberately so. This thesis aimed to delineate the space of possible model behaviour and in that way get some understanding of what we can expect to learn from the models. Of course, the space is restricted to 'relevant' models, interesting enough to have been studied for several decades. Those, by and large, come in two flavours: the horizontal naming games and vertical iterated learning tradition. The main contribution of the thesis is an attempt to unify these two traditions in a new model: the Bayesian language game.

BAYESIAN NAMING GAME   This Bayesian language game is a simple extension of the Bayesian *naming* game, also proposed in this thesis. In the Bayesian naming game, all agents haven an internal language θ from which they draw words in every encounter. After observing utterances, they update their beliefs using Bayesian updating. Concretely, we studied a Dirichlet-categorical variant of this framework, meaning that the prior beliefs of the agent are modelled by a Dirichlet distribution. Although mathematically simple, the game exhibits a rich behaviour. First of all, the population converges to a shared, stable language **??**. Second, this language *reflects the bias*: it is clearly shaped by both the innate biases and the cultural process, in a non-trivial way. That means that different lineages develop different languages **??** Third, the game develops in three stages, metaphorically called 'infancy', 'puberty' and 'adulthood'. In the first two stages the language primarily develops its characteristic, contingent shape, and during 'adulthood' it stabilises.

The Bayesian naming game addresses the lack of language stability in the Bayesian iterated learning models, but it also incorporates one of the main innovations of the Bayesian models: the explicit representation of innate biases **??**. Further, in this game the innate biases can be transparently separated from past linguistic experience. Consequently, the 'total' beliefs of an agent are, quite literally, the sum of innate biases and past experience. Another innovation of the Bayesian iterated learning models was the study of different strategies agents use for selecting languages: using the most likely language under the posterior (MAP or maximising) or sampling a language. This can

A. Iterated learning ($\gamma = 1$)

B. Naming Game ($\gamma = \infty$)

C. Quick turnover ($\gamma = 10$)

D. Medium turnover ($\gamma = 100$)

S–S   M–S   S–M   M–M

—•— Aggregate language $\bar{\pi}$ —— Expected language π ▬ External language ψ

be directly translated to the Bayesian naming game in the form of a parameter η, together with a production strategy, parametrised by ζ, derived from the naming game literature. After all, the game is primarily a *naming* game and was accordingly shown to implement a kind of lateral inhibition, one of the alignment strategies used in that field.

BAYESIAN LANGUAGE GAME   The Bayesian *language* game extended the Bayesian *naming* game by changing the population structure so that it can either take the form of a naming game, or the form of an iterated learning model. The language game performs a random walk through a population of fixed size. It moreover adds a life expectancy, and when an agent dies it is replaced by a newborn agent. The random walk, when unraveled, forms a chain (iterated learning), but since it is random, it simultaneously approximates homogeneous mixing (naming game). The crucial parameter that interpolates between naming games and iterated learning is the life expectancy γ. The Bayesian language game is therefore parametrised by three crucial parameters: η for the language strategy, ζ for the production strategy and γ for the life expectancy. This, then, is the 'space of agent-based models' whose behaviour we aimed to characterise: the parameter space (η, ζ, γ) of the Bayesian language game.

The eight extreme cases, with $\eta, \zeta, \gamma \in \{1, \infty\}$ are of special interest. With $\gamma = 1$ the language game reduces to iterated learning, and with $\gamma = \infty$ it reduces to a naming game. The pure sampling strategy (sample–sample) corresponds to $\eta = \zeta = 1$; the mixed sampling-maximising strategies to $\eta = 1, \zeta = \infty$ (sample–MAP) and $\eta = \infty, \zeta = 1$ (MAP–sample); and the pure maximising strategy to $\eta = \zeta = \infty$ (MAP–MAP). A systematic search through the space suggested that the behaviour of the game would relatively smoothly change between these eight extreme cases. Characterising their behaviour therefore goes a long way to charting the behaviour one can expect from agent-based language models. The result of cultural evolution in the extreme

cases (and two intermediate population turnover rates) are shown again in figure 1. The main conclusions it suggests are:

- Bayesian naming games results in languages reflecting the bias, but shaped by cultural evolution.
- In iterated learning models, no agent faithfully represents the external language.
- Pure sampling strategies stay close to the bias, either perfectly (iterated learning), or imperfectly (naming game)
- Mixed strategies differently 'exaggerate' the bias: maximising languages results in some kind of pruning, maximising productions in some kind of exponentiation.
- Pure maximising strategies result in degenerate languages.
- The language spoken in an iterated learning model seems predictably determined by the biases — even for maximising strategies.
- Intermediate life expectancies interpolate between these two extremes.[1]

That, in short, that was the answer this thesis formulated on the first subquestion addressed in this thesis: what kind of behaviour can we expect from agent-based models of language evolution? Well, see figure 1, and interpolate!

EMPIRICAL VALIDITY   That might be a good start, but what does it tell us about language evolution? In all fairness, I do not think I have formulated anything close to a clear answer to this question. Perhaps because this is a not exactly an easy question.

In the second part of the thesis I identified numeral systems as an interesting test case for models of language evolution. I gave various reasons, pertaining to the amount of linguistic data available, the sheer size of the design space (which makes for an interesting search problem), the balance between expressivity and simplicity in numerals and, most importantly, that the structure of numerals themselves suggest a reconstruction of their origins. The last chapter tried to connect that reconstruction to agent-based models of language evolution. That led to some interesting findings, such as the fact that the biases of agents and the constraints of the domain interact in different ways. Or the sensitivity of naming games to subtle frequency biases in the underlying domain.

In fact, that finding is perfectly illustrative of the problem: all agents seem to be doing in these models is a more or less sophisticated form of frequency administration. Administration surely helped the rise of human civilisation, but perhaps not this kind of administration. *Perhaps*. The point is that adapting to frequencies describes a *mechanism* that amounts to an explanation of the observed phenomenon — but is it the right one, is it justified? Frequency 'administration' could explain the standardisation of bases, as we have seen in the last chapter. But one could also explain that by using arguments of communicative efficiency, as Hurford (1987) also proposed. Which one to prefer? Or take a different example: the idea that compositionality emerged because it balances expressivity and compressibility. The explanations suggested for one of the purest compositional structures, that of numerals, point in a quite different direction: a counting practice, grouping objects, or, later, processes of grammaticalisation. Which one is it?

As soon as one leaves the world of models and enters the world of language, in this case numeral systems, it becomes clear that you have moved to a different, lower level of description. One could conclude that the higher level description is 'wrong', but that

[1] With one notable exception: there might be stable language change in between, but possibly only for a very limited parameter range.

would be too simple. Descriptions at different levels are rarely perfectly aligned — that's the whole point of having different levels of descriptions in the first place. But what, then, is the conclusion? Perhaps that it is time to see if the levels of description align at all. That is, to see if the models can explain actual linguistic data, outside the computer or the lab.

I did not like Berwick and Chomsky's paper, as the reader could no doubt tell from the introduction. Just as pretty much any other field of science, the 'Kirby-type work' has problems and good, sharp critiques are refreshing. But the style — demeaning, territorial, grotesque. I mean, the academic equivalences thereof. Science, the knowledge commons, those are, in the end, beautiful, praiseworthy things and you would hope for a more more communal spirit.[2] In any case, there is one point at which I could not disagree with Berwick and Chomsky (2017): when they write that, as far as they know, the "whisper down the lane" properties (slang for 'iterated learning') have not been applied to "the empirical phenomena of language change that linguists have actually measured" outside the lab or simulations. It is just that I would prefer not to read that as a fatal blow, but as a programme.

I will leave it at that. The remainder of this chapter lists the main contributions, some points of discussion and future work.

Main contributions

The main two contributions of this thesis are (1) the Bayesian Naming Game, an attempt at unifying the naming game with the Bayesian model of iterated learning, and (2) the identification of numeral systems as a testbed for models of cultural language evolution. Various other additions to existing literature have been made throughout the thesis. The ones I believe to be of most interest are listed below.

- **Lineage-specific languages can reflect a prior.** A demonstration that longer lifespans can lead to lineage-specific, shared languages that reflect the prior, but do not mirror it exactly. This suggests that cultural evolution *can* change languages in nontrivial ways, while still being subject to innate constrains.
- **Random walks as a transmission model.** A new model of cultural transmission in the form of a random walk through a population of fixed size. This combines the transmission chains with homogeneous mixing.
- **Connecting naming games to iterated learning** The thesis tried to connect two agent based modelling paradigm. Most specifically this results in a direct mathematical parallel between the model of De Vylder and Tuyls (2006) and Bayesian iterated learning models.
- **Numerals as a testbed for language evolution.** I have argued that numeral systems are a promising test case to relate models of language evolution with actual language. Numeral systems (i) allow one to model a subsystem of natural language directly; (ii) there is plenty of linguistic variation (i.e. the search of possible systems); (iii) the cognitive capacity for numerosity has been studied extensively;

and (iv) there are benchmarks in the form of a reasonable reconstruction of the origins of numeral systems, which is moreover supplemented by recent empirical studies.

- **Realistic population turnover using Weibull distributions.** Proposed a more realistic death-process for iterated learning models based on the Weibull distribution, often used to model life-expectancy. Earlier iterated learning studies have often assumed constant hazard rates, which are unrealistic.
- **Revisited Hurford's base games.** I have shown that Hurfords early simulations of the standardisation of a base are variants of the Naming Game.
- **Domain-adaptivity in the Base Game.** Demonstrated that the Base Game adapts to the domain in the sense that the behaviour depends on frequency patterns determined by domain.
- **Counting Game.** A new type of naming game in which agents negotiate a counting sequence. In this model simple numerals derive their meaning solely from their position in the sequence.
- **Packing strategy without grammar.** A reformulation of the packing strategy independent of the phrase-structure grammar proposed in Hurford (1975) in terms of the semantic (arithmetic) structure of the numerals. This is particularly relevant for empirical studies of this proposed universal.

Future work

Throughout this thesis, possibilities for future work have been identified. Let me list the most practical points first, and then move on to the more interesting questions.

- **Other lateral inhibition strategies.** This thesis only compared Bayesian updating has only to the *basic* lateral inhibition strategy in Wellens (2012). Wellens also mentions the so-called *interpolated* LI strategy, which can be interpreted as the Rescorla-Wagner or Widrow-Hoff rule. In hindsight, that seems closer to Bayesian updating (in form, at least) and future work could address the exact relation.
- **Measures for the Bayesian naming game.** The measures used to analyse the Bayesian naming game could be improved by developing appropriate variants of measures used in naming games (communicative success, number of (unique)). The methodology of chapter **??** could be a starting point.
- **Scaling relations.** The change in convergence time in the Bayesian naming game under varying population, vocabulary and bottleneck size could be investigated more thoroughly, either empirically or analytically.
- **Models of life-expectancy.** Chapter **??** argued for more realistic models of life-expectancy without actually analysing whether they make a difference in the simulations. A systematic analysis would be valuable.
- **Random walks vs. homogeneous mixing.** Preliminary experiments did not reveal any systematic differences between random walks and homogeneous mixing in the Bayesian naming game, but this should be checked in a principled fashion.
- **Mathematical problems.** The mathematical analyses gave rise to several prob-

lems. The first concerns the characterisation of the posterior distribution of an agent with a maximising word strategy $\zeta > 1$; see appendix **??**. Another concerns the bias in Hurford's additive base game: has this distribution (proportional to the difference of two harmonic numbers) been studied before?

OTHER BAYESIAN NAMING GAMES   This thesis mainly developed the Dirichlet-categorical naming game as the simplest instantiation of a more general framework. Future work could study other models inside the same framework. An interesting possibility would be to drop the conjugacy assumption, and consider more general probabilistic models. In fact, the Dirichlet-categorical model surfaced from early experiments with Bayesian agents represented as so called *probabilistic programs*. Probabilistic programming (see Ghahramani 2015 for an overview) provides a remarkable representational flexibility by using programming languages to specify the models, and their universal approximate inference algorithms (sampling or variational methods) significantly speed up development time. The downside in our context would be that poor inference could have unexpected results when repeated for tens of thousands of rounds and thus trouble the analyses.

CONNECTIONS TO BIOLOGICAL EVOLUTION   Reali and Griffiths (2010) were the first to connect Bayesian models of iterated learning to biological evolution, i.e. the Wright-Fisher model. Since the Bayesian Naming Game is in many ways identical to their model, it seems worthwhile to explore further connections with models developed in population genetics.

NONPARAMETRIC EXTENSION OF THE BAYESIAN NAMING GAME   The Bayesian Naming Game made the simplifying assumption that the number of words in the population is fixed, an assumption made earlier in De Vylder and Tuyls (2006). The model would stay closer to the original naming games if this assumption could be dropped. Invention after all plays a crucial role in the Naming Game (Steels 2011), but more generally it has been argued that inventors play an important role in language change and evolution (Hurford 1987). Fortunately, there is a straightforward extension of the Bayesian Naming Game which does not fix the number of categories — is not *parametric* in that sense. This model would use an infinite analogue of the Dirichlet distribution, a Chinese restaurant process, that allows for the use of an arbitrary number of words. Every agent then faces the choice of either using an old word or inventing a new one. This extension has been discussed in the iterated learning literature before Reali and Griffiths (2010) and Burkett and Griffiths (2010).

Combined with population turnover, the non-parametric model seems to solve another issue: that words never disappear from the population completely. In classical lateral inhibition games, words scoring below a certain threshold are removed. This mechanism allows the emergence of *efficient* vocabularies. Population turnover could replace that mechanism in a non-parametric variant of the Naming Game: low-frequency words might simply not be acquired by newborn agents, and consequently die out. If this prediction is indeed true, it suggests a very exciting possibility. Suppose one splits the population in two parts at a certain point in time, and let both halves continue the game separately. Given the stochasticity of the game, it seems likely that dif-

ferent words would die out, effectively leading to different lineages. Future work could investigate such 'linguistic speciation' in a non-parametric extension of the Bayesian Naming Game.

Note that this also highlights a problem with the notion of 'lineage specificity' as I have used it. 'Lineages' were understood to mean 'runs', since it has not been shown that one can, say, split the population and develop two distinct branches. For the Bayesian naming game, the later seems moreover unlikely, since the language develops its characteristic shape primarily in the early phase of the game, before agents reach 'adulthood'. After that, the population only converges to a completely stable language, which is not realistic either (Kirby 2001). Increasing population turnover addresses this, and the nonparametric suggestion outlined above might solve the former.

MATHEMATICAL ANALYSIS OF THE BAYESIAN NAMING GAME   Simulations indicate that coherence emerges in the Bayesian Naming Game, but I have not provided a formal proof of this. In fact, there are two separate questions: (i) do all agents converge to the same distribution? And (ii) can one characterise this distribution in terms of the parameters η, ζ and γ? It seems unlikely that this is a simple problem, since even the special case of the naming game has so far resisted analytical scrutiny (De Vylder and Tuyls (2006) only provide analytical results for a deterministic variant of the game). One possible line of attack would investigate whether the limiting distribution is indeed a sample from some distribution around the prior. If so, a characterization of that distribution would be likely to provide valuable further insights into many models of cultural language evolution simultaneously.

And the list continues. But that's it for now.