Work & Play: Simulating language contact

Roland Mühlenbernd & Jason Quinley, University of Tübingen

Introduction. English shares cognates with many other languages, notably German and French. The Norman conquest of 1066 infused new words into the English lexicon, creating a divergence in register and context. Historians point to social conditions like the lifestyles of the Anglo-Saxon laborers and the French-speaking Norman nobility as the cause. But could the results have been different?

Contact and Context. The distinctions between animal and meat (as seen in Table 1) follow a salient trend in the English language: vocabulary of French origin tends to the prestigious, whereas words of Germanic stock often fall into working class topics; i.e. the laboring class has to work (Ger. *Werk*), the upper class can play (from Fr. *plaisir* = pleasure). The *partial blocking* seen in English suggests that i) words like *swine, cow* and *sheep* once bore an animal/meat polysemy like their German cognates and were used context-dependently and ii) lost this polysemy by adopting French alternatives, causing the partial blocking. We call this *emancipation from context dependence*.

We simulate how language contact impacts usage by applying signaling games (Lewis, 1969) to social networks. We investigate i) the emancipation from context dependence and ii) the social parameters responsible for the salient systematic division of meaning between words of Germanic and French origin.

animal		meat	
swine	(Ger. Schwein)	pork	(Fr. porc)
cow	(Ger. Kuh)	beef	(Fr. bœuf)
sheep	(Ger. Schaf)	mutton	(Fr. mouton)
deer	(Ger. Tier)	venison	(Fr. venaison)

personal	abstract
freedom	liberty
knowledge	science
belief	faith
brotherly	fraternal

Table 1: Systematic division of meaning space between words of Germanic and French origin. **Left:** division betweens words for animals and their meat. **Left:** another example of such a division between concrete (Germanic origin) and abstract (French origin) concepts.

Experimental Set-Up. Language use is modeled by the repeated *context signaling game* (CSG), a Lewisean signaling game, played between two agents (speaker S and recipient R) and augmented by both a set of contexts $C = \{\text{dinner, farm}\}$ and probabilities on these contexts based on social status $\sigma \in [1, 9] \subseteq \mathbb{N}$, assigned to each agent. Nature N chooses one of those contexts depending on the speaker's social status, and the sender attempts to communicate an information state from a set $T = \{'meat', 'animal'\}$ by using one of some possible messages from a set $M \subseteq \{\text{"swine", "porc"}\}$. If the receiver construes the message with the right interpretation state, communication is successful. Figure 1 depicts the game tree of such a CSG, where (un-)successful communication is denotes at the leaves with 1 or 0, respectively.

300 agents with a randomly chosen social status and arranged on a *scale-free* network (Jackson, 2008) play the CSG repeatedly as both speaker and recipient with their neighbors. Based on previous encounters, agents learn a communication strategy via a *belief learning* dynamic (e.g. Nachbar, 2008).

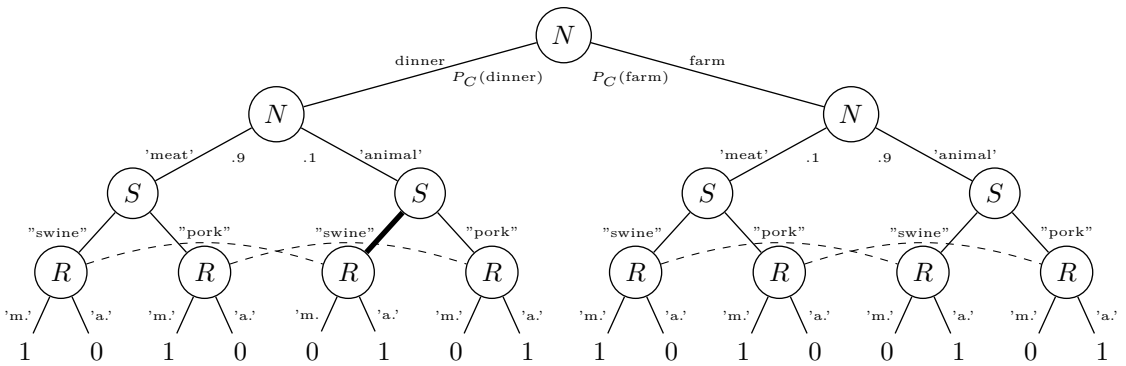

Figure 1: Game tree depicting the decision process of the Sender and Receiver in the CSG. Nodes connected with a dotted line are indistinguishable to the Receiver. Bold edges indicate strategy S_1.

At a simulation's start the agents' lexicons contain only the expression "swine". We then simulate the conquest of 1066 by replacing 10% of the agents by Norman invaders, agents whose lexicons contain only the expression "pork". When an agent as a recipient encounters an unknown expression, he adopts it to his lexicon. The simulation ends when every agent's lexicon contains "swine" and "pork" and a unique strategy governs the whole society (that happened in 100% of all simulation runs).

Result 1: Context-Dependent Language Use. The initial network's agents learn a context-dependent signaling strategy, learning that the topic 'meat' is more probable in the dinner context and topic 'animal' in the farm context when they interpret the expression "swine".

Result 2: Emancipation from Context Dependence. The invasion, done by replacing 10 randomly chosen agents with invaders, triggers the the word "pork" spreading, entering every agent's lexicon. Every agent, as both speaker and as recipient, learns a new *context-independent* strategy). The new expression makes the context irrelevant. After each trial every agent has learned one of two possible strategies: S_1 or S_2 (Figure 2 (left)). Hundreds of simulations revealed that i) in each trial only one strategy spreads and stabilizes society-wide and ii) both strategies' emergence is equiprobable. Since there is nothing, for example, about the word *pork* that makes it mean *meat*, we see that the emergence of the expected strategy S_1 for using the word of French origin for the meat and word of Germanic origin for the animal is as probable as the other way around. But if chance determines whether S_1 or S_2 emerges, then how can we explain that language use like S_1 is predominant?

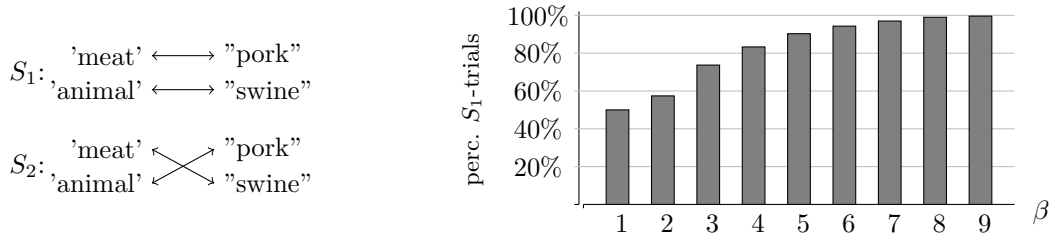

Figure 2: *Left:* Strategies S_1 and S_2 emerge equiprobably if agents are replaced randomly. *Right:* The higher the minimum status (β) for replacement, the more probable the expected strategy S_1.

Result 3: Influence of Social Status and Structure. In the previous simulation runs, randomly chosen agents were replaced by invaders. We now account for the fact that the Normans probably occupied high social positions in two ways: first, we consider the social status of the replaced agents as $Pr(C_D)$. I.e. higher status agents are more likely to be in a *Dinner* than *Farm* context. Figure 2 (right) depicts the percentage of trials resulting with the expected strategy for each baseline of social status for the invaders. We see that the higher the social status of agents replaced by invaders, the higher the probability for the emergence of the expected strategy. Second, in accordance with results on wealth distribution in societies (Jackson, 2008), we interpret wealth as number of business (or speaking) partners. This gives a rationale for replacing the hubs in the original society with Norman invaders. This additional parameter accelerated the probability for the adoption of the expected strategy.

Discussion. We showed that agents in a social network playing the CSG with a single-word lexicon resolve polysemy context-dependently. The simulated Norman invasion provided lexical alternatives, allowing speakers to distinguish between previously context-dependent meanings. But without further assumptions, the new words could have described any meaning equiprobably. By considering social status, the probability shifted to the expected strategy. In the full paper, we detail further effects of network structure, costly signaling, and social status on the agents' communication strategies.

References. Blumenthal & Kahane (1979), *Decline and Survival of Western Prestige Languages.* ⋆ Jackson (2008), *Social and Economic Networks.* ⋆ Lewis (1969), *Convention.* ⋆ Nachbar (2008), "Learning and Evolution in Games: Belief Learning", *The New Palgrave Dictionary of Economics.*