Why Synonymy is Rare: Fitness is in the Speaker

James R Hurford

Language Evolution and Computation Research Unit,
School of Philosophy, Psychology and Language Sciences,
University of Edinburgh, Edinburgh EH8 9LL, Scotland, UK,
jim@ling.ed.ac.uk,
http://www.ling.ed.ac.uk/ jim/

Abstract. Pure synonymy is rare. By contrast, homonymy is common in languages. Human avoidance of synonymy is plausibly innate, as theorists of differing persuasions have claimed. Innate dispositions to synonymy and homonymy are modelled here, in relation to alternative roles of speaking and hearing in determining fitness.
In the computer model, linguistic signs are acquired via different genetically determined strategies, variously (in)tolerant to synonymy or homonymy. The model defines communicative success as the probability of a speaker getting a message across to a hearer; interpretive success is the probability of a hearer correctly interpreting a speaker's signal. Communicative and interpretive success are compared as bases for reproductive fitness. When communicative success is the basis for fitness, a genotype evolves which is averse to synonymy, while tolerating homonymy. Conversely, when interpretive success is the basis for fitness, a genotype evolves which is averse to homonymy, while tolerating synonymy.

1 Synonymy and Homonymy in Languages

Synonymy, at least of the pure variety, is rare. If found at all, it is usually in

- contact between languages (*napkin/serviette, eggplant/aubergine*) or dialects (*flat/apartment, bonnet/hood, boot/trunk, mobile/cellphone*),
- handy abbreviation (*microphone/mike, bicycle/bike, hippopotamus/hippo*),
- specialized domains where euphemism is rife, such as sex (*fuck/shag/* ...), death (*croak/expire/* ...) and bodily functions (*shit/crap/* ...).

Human avoidance of synonymy is plausibly innate; theorists of starkly differing persuasions have claimed so. In the functionalist/empiricist tradition, Clark's Principle of Contrast [2] attributes synonymy-avoidance to a (presumably inborn) "pragmatic" tendency of humans to seek and/or create new meanings, rather than accept one meaning for several different forms. Markman [7] notes a disposition in children to assume, initially at least, that no two words may overlap in meaning. The formalist Wexler has proposed an innate Uniqueness

Principle [16], which prevents the child from internalizing more than one form per meaning; Pinker [9] also subscribes to this Uniqueness Principle.

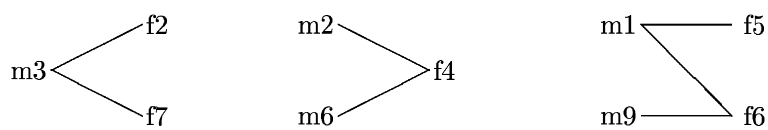
By contrast, homonymy is common in languages. Consider the many and unrelated senses of such words as *bank, list, file, toast* and *port.* A priori, this is puzzling, as one would suppose a well-designed language to be able to tolerate synonymy without threat to its communicative efficiency, but to outlaw homonymy, as it potentially gives rise to interpretations by a hearer differing from the speaker's intended message. Computer languages and command systems, for example, frequently allow aliases (alias synonyms), but often cannot handle different intended uses of the same term in the same syntactic context (homonymy).

This paper suggests one reason why synonymy-avoidance is in some sense innate, and why there is no parallel innate tendency to avoid homonymy.

2 The Model

2.1 Initial Vocabularies

Each simulation begins with an initial vocabulary of 50 signs. A sign is a form-meaning pair. To compose a single sign, a form is drawn randomly from a set of 100 possible forms ($\mathbf{f}_1 \ldots \mathbf{f}_{100}$), and a meaning is drawn randomly from a set of 100 possible meanings ($\mathbf{m}_1 \ldots \mathbf{m}_{100}$). This method of composing an initial vocabulary gives rise to accidental instances of both synonymy and homonymy, as illustrated below.



Synonymy (two signs) | Homonymy (two signs) | Synonymy and homonymy (three signs)

In the initial 50-sign vocabularies there were on average between eight and nine such instances of synonymy and of homonymy.

2.2 Population of Learners

All simulations started with populations of 30 agents, composed of 10 innate 'synonym-rejectors', 10 innate 'homonym-rejectors', and 10 innate 'allowers'. These distinct genotypes competed over many generations, until one genotype had completely eliminated the others. The distinct genotypes are defined by their learning behaviour when exposed to a vocabulary of signs. Learning proceeded by each learner being exposed to each sign in the vocabulary of the preceding generation. As the name suggests, when a synonym-rejector was exposed to several distinct signs involving the same meaning (but different forms), it only acquired the first, ignoring all the others. Likewise, when a homonym-rejector was exposed to two or more signs involving the same form (but different meanings), it only acquired the first, ignoring the others. An agent with the 'allower' strategy was more permissive, acquiring all the signs to which it was exposed.

After all the initial population had acquired their individual vocabularies, each individual's fitness was assessed on the basis of how well it could transmit messages to, or interpret signals from, the rest of the population.

Where an agent has several (f) forms for one meaning (i.e. synonyms), it chooses randomly between those forms when attempting to convey that meaning. Thus with f forms for one meaning, the probability of using any particular form for that meaning is $1/f$.

An agent's interpretation of a signal was modelled under three 'context' conditions, 'zero-context', 'full-context' and 'half-context'. In the zero-context condition, where a hearer agent has several (m) meanings associated with one form (i.e. in a case of homonymy), it chooses randomly between those meanings when it hears that form. Thus if a hearer has m different meanings for one form, the probability of his interpreting that form as any particular meaning is $1/m$. In the full-context condition, a hearer agent is always empowered to know which meaning of a form the speaker agent intended; so in this condition the probability of interpreting a form as the meaning intended by the speaker is 1, and the probability of interpreting it as any other meaning is 0. In the half-context condition, a hearer agent is partly biased toward the meaning of a form which a speaker agent intended. This is precisely defined so that the probability of the hearer interpreting a form as the meaning intended by the speaker is $(1+1/m)/2$, where the form has m meanings in the hearer's vocabulary at the time.

Two measures of fitness were defined, Communicative Potential (CP) and Interpretive Potential (IP), as follows. For any pair of agents, S and H, S's communicative potential relative to H and relative to a given meaning M is expressed as $CP(S,H,M)$ and defined thus:

$$CP(S,H,M) = \sum_{i=1}^{n}[P(S \text{ transmits form } i \text{ for } M) \times P(H \text{ interprets } i \text{ as } M)]$$

where n is the number of forms associated with M in S's vocabulary. Thus $CP(S,H,M)$ is the probability, given that S intends to convey M, of S successfully communicating M in an encounter with H. The converse of CP is IP (interpretive potential). $IP(H,S,M) = CP(S,H,M)$. $IP(H,S,M)$ is the probability, given that S is thinking of M, and transmits a form for it, of H interpreting that form as meaning M. A more general measure $CP(S,H)$ can be defined, not relative to any particular meaning, but averaging over all meanings in the system.

$$CP(S,H) = \frac{\sum_{i=1}^{n} CP(S,H,i)}{n}$$

where n is the number of meanings in the system. $CP(S,H)$ is the probability, given that S intends any arbitrary meaning, of S successfully conveying that meaning in an encounter with H. Conversely, $IP(H,S) = CP(S,H)$. $IP(H,S)$ is the probability of H successfully interpreting any form emitted by S. An individual's CP (or IP) relative to the whole population is obtained by averaging over its CP (or IP) relative to all other individuals in the whole population. (These are exactly the same fitness measures as used in [3] and [8].)

2.3 The Simulation Cycle

Each simulation went repeatedly through the following cycle:

1. Establish CP (and/or IP) for each individual relative to the whole population,
2. Nominate parents, on weighted basis of CP (and/or IP),
3. Breed new population of 30 individuals from nominated parents, passing on a parent's innate acquisition strategy ('allow', 'syn-rej', or 'hom-rej'),
4. Each new individual acquires a sign-set on the basis of its inherited innate acquisition strategy and exposure to the complete sign-set of the previous generation of the population,
5. If there is still variation in the population, go back to step 1 and re-cycle.

Two orthogonal sets of conditions were investigated. One set of conditions was the 'context effect', mentioned above; here there were three conditions, zero, half, and full. The other set of conditions was the fitness measure; here again there were three conditions. In one condition, an individual's fitness was proportional to its CP, in another, to its IP, and in the third condition to the average of its CP and IP. Thus fitness was variously identified with success as a speaker, success as a hearer, and overall success as both a speaker and hearer. The two intersecting sets of conditions gave a 3×3 experimental design. For each of the 9 combinations of conditions, 100 simulation runs were carried out. A run terminated when the whole population was homogeneously of one innate type of sign-acquirer. Each of the 900 runs started with a different randomly generated initial vocabulary, as described in section 2.1.

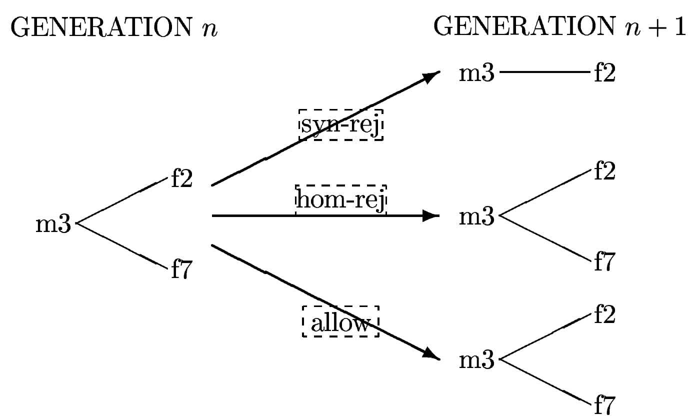
2.4 Results

Figures in cells in the table below show how many simulations, out of 100, each genotype 'won'.

	CONTEXT EFFECT					
	zero		*half*		*full*	
$fitness = CP$	*allow*	6	*allow*	20	*allow*	48
	syn − rej	94	*syn − rej*	80	*syn − rej*	52
	hom − rej	0	*hom − rej*	0	*hom − rej*	0
$fitness = IP$	*allow*	0	*allow*	12	*allow*	32
	syn − rej	4	*syn − rej*	8	*syn − rej*	39*
	hom − rej	96	*hom − rej*	80	*hom − rej*	28
$fitness = \frac{CP+IP}{2}$	*allow*	10	*allow*	29	*allow*	47
	syn − rej	41	*syn − rej*	59	*syn − rej*	52
	hom − rej	49	*hom − rej*	12	*hom − rej*	1

Dissection of the Basic Result The most striking results appear in the top two cells of the left-hand column. These directly compare selection for communicative potential with selection for interpretive potential, in the zero-context condition. In the zero-context condition, no help is given to hearers in disambiguating ambiguous signals; they choose at random from the possible meanings of a form. These boxes show clearly that, where fitness is associated with success in getting ones meaning across (i.e. with CP), selection favours individuals who resist acquiring synonyms, while agents who resist acquiring homonyms are never favoured. Conversely, where fitness is associated with success in correctly identifying another agent's meaning (i.e. with IP), selection favours individuals who resist acquiring homonyms, while agents who resist acquiring synonyms are very seldom favoured.

If this result is puzzling, it is helpful to consider a single small example as a microcosm of what happens in the simulation. Consider the case of learners exposed to two signs, exhibiting synonymy. Exposed to such data, synonym rejectors acquire a single sign, whereas homonym rejectors and allowers acquire both signs, as diagrammed below.

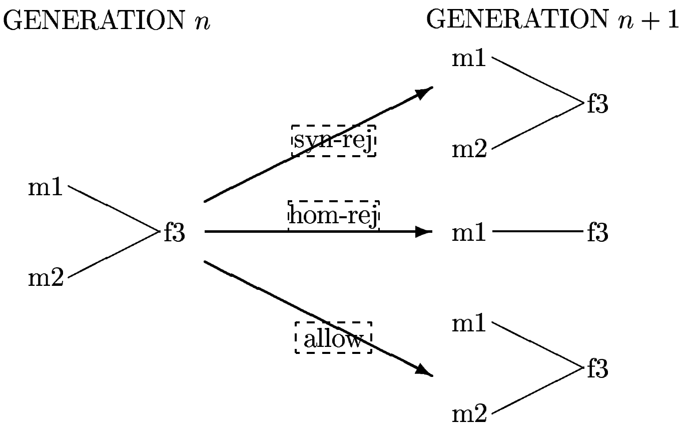


Now consider a mini-population consisting of just three agents, one of each type, with minimal vocabularies as in this diagram. That is, the synonym rejector has a single sign relating meaning $\mathbf{m}_3$ to form $\mathbf{f}_2$, whereas the homonym rejector and the allower both have two signs involving this meaning, relating it to forms $\mathbf{f}_2$ and $\mathbf{f}_7$. If these three agents now attempt to communicate with each other about meaning $\mathbf{m}_3$, their rates of success (i.e. their CP and IP relative to meaning $\mathbf{m}_3$) are as summarized in the following table.

	syn-rej	hom-rej	allow	CP
syn-rej	1.0	1.0	1.0	1.0
hom-rej	0.5	1.0	1.0	0.833
allow	0.5	1.0	1.0	0.833
IP	0.667	1.0	1.0	

The marginal figures here are averages[1]. It can be seen that, in this case, synonym rejectors make themselves harder for the other types to communicate with, thus disadvantaging them, while keeping their own perfect communicative potential (CP). But the synonym rejector's IP is less than that of the other two types.

Next consider a converse case of learners exposed to two signs exhibiting homonymy. Exposed to such data, homonym rejectors acquire one sign, whereas synonym rejectors and allowers acquire two, as diagrammed below.



Again, we can imagine a mini-population of just three agents, one of each type, with the mini-vocabularies as on the right of this diagram. If these agents try to communicate with each other about meanings $\mathbf{m}_1$ and $\mathbf{m}_2$, they will have varying success, as shown in the following table. (Here, there are two numbers in the main cells, the first relating to communication about meaning $\mathbf{m}_1$ and the second to communication about meaning $\mathbf{m}_2$. Again, the marginal CP and IP figures are averages, and the earlier footnote still applies.)

	syn-rej		hom-rej		allow		CP
syn-rej	0.5	0.5	1.0	0.0	0.5	0.5	0.5
hom-rej	0.5	0.0	1.0	0.0	0.5	0.0	0.333
allow	0.5	0.5	1.0	0.0	0.5	0.5	0.5
IP	0.417		0.5		0.417		

It seems reasonable to interpret the homonym rejector as being unable either to transmit or receive messages involving meaning $\mathbf{m}_2$, thus getting a zero score for both CP and IP relative to that meaning. It can be seen that all types have a problem with homonyms; the marginal averages are lower in this table than in the preceding table. But, on IP score, the homonym rejector does better than the other types. Conversely, on CP score, the synonym rejector and allower outperform the homonym rejector.

[1] These are averages over all figures in the rows and columns, thus including the case where an agent talks to one of its own type, as well as to agents of the other two types. If, alternatively, the figures in the main diagonal of the table are ignored, the resulting averages make the central point of this subsection even more strongly.

Putting the two tables together, thus combining the two possible basic scenarios (synonymy and homonymy), it is clear that the synonym rejector outperforms the other two types on CP, and the homonym rejector outperforms the other two types on IP. This dissection of the typical elementary interactions in the simulations shows why, in the overall results at the beginning of this section, in the zero-context condition, synonym rejectors were by far the biggest winners when fitness was correlated with CP, and conversely why homonym rejectors were easily the biggest winners when fitness was correlated with IP.

Other Results

The Context Effect. The context effect modelled the influence of context in helping to disambiguate homonyms. The full context condition effectively banished the problem of homonymy, and the half context condition partially banished it. The homonym rejection strategy is an innate solution to the homonymy problem, shown to be effective by the results in the middle box of the left-hand column of the main results table. Banish the problem, and there is no need for a solution. Thus, as one moves rightward across the table, in the middle row, homonym rejectors fare progressively worse. (In the top row, they could not get any worse, starting and staying at zero wins against the other strategies.)

The asterisk in the middle right-hand cell of the main results table above indicates that the numbers in this cell only add up to 99, as one of the runs did not terminate in the time allotted. The numbers in this cell are more finely balanced between all three competing strategies than in any of the other cells. In this combination of conditions, ‘full-context’ gives all agents a form of mind-reading to disambiguate ambiguous signals, and fitness is based on IP. Thus these conditions give no sign-acquisition strategy a clear advantage.

Fitness Averaged between CP and IP. The bottom row of the main results table shows numbers of winners of each type when fitness was based on the average of CP and IP (with various context effects). Here, not surprisingly, with zero context effect (left-hand box), synonym rejectors and homonym rejectors were about evenly balanced, but both had a clear advantage over allowers. Now look rightward in the table along the bottom row, first through half-context diminishing the problem of homonymy, and then to full-context eliminating the problem entirely; homonym rejectors become progressively and markedly less successful, with their niche being taken over largely by allowers and to a lesser extent by synonym rejectors.

3 Discussion

3.1 Relation to Other Work

The model used in this study bears obvious similarities to several other models, but differs in various ways from all of them.

In the present model, the signs are learned, and not innate like those in the models of [15, 1].

The present model is not a model of how populations, starting with no shared conventional meaning-form pairings, evolve socially coordinated signalling systems. In the present model, the same initial (randomly generated) vocabulary is given to each member of the population in the initial generation; agents of different innate types subsequently modify these vocabularies, but only by eliminating some signs. In this way, the present model differs significantly from those in [8, 10, 11, 13, 14].

In K.Smith's model [10], various innate learning biases compete over many generations, as in the present model. He found that suitably biased (homonymy- and synonymy-avoiding) innate learning rules did not evolve in the absence of a pre-existing shared vocabulary in the population. In another model [11], K.Smith showed that only certain innate learning biases can give rise to the cultural evolution, over many generations, of a communally shared set of meaning-form pairings. Combined with the results obtained in the present model, these studies give us a chicken-and-egg problem. This model shows that innate anti-synonymy and anti-homonymy biases may evolve naturally, given a pre-existing shared code; K.Smith's studies show that a common code cannot arise unless the innate biases are already in place. In a further study, [12] he argues that the required biases may have arisen for reasons independent of communication. It would be worthwhile, in future work, to try to solve the chicken-and-egg problem here with a more direct approach, involving 'mutual bootstrapping' of the cultural evolution of a code and the biological evolution of the learning biases. This might be possible by starting with very small sets of meanings and available signals, possibly as small as 2. All the studies reported here started with larger sets of available meanings and forms, e.g. 10 meanings in K.Smith's case, and 50 meaning-form pairs in the present study.

The present model is technically an instance of the 'Iterated Learning Model' (ILM) [4, 6, 11]. Such models explore the evolution of signalling systems under conditions of repeated cultural transmission over many generations. In the basic form of the ILM, there is no biological evolution; all the interesting changes are in the culturally transmitted abstract system symbiotically 'inhabiting' the population. An extended version of the ILM is the 'Evolutionary Iterated Learning Model' (EILM) in which the coevolutionary interactions between biological and cultural evolution can be explored; an instance is [5]. In the present model, there is indeed both cultural transmission (of signs) across generations and biological transmission (of sign-acquisition strategies), and so the model is technically an instance of an EILM. But in the current model, the dynamics on which I have focussed exhibit neither interesting changes in the culturally transmitted vocabularies nor coevolutionary interaction between cultural and biological mechanisms. The interesting conclusions relate to the evolved genotypes, under the various conditions. In this sense, the present model is more like a genetic algorithm in spirit, with experimental variation of the fitness function.

There is a rather obvious, but trivial, sense in which the culturally transmitted vocabularies evolve in the present simulations. When a population is totally taken over by innate synonym-rejectors, the population's shared vocabulary, by definition, contains no synonyms. Likewise, the vocabulary of a population taken over by homonym-rejectors contains homonyms. Further experiments, not reported in detail here, show that, if random mutation between the three genotypes is permitted in a population which has culturally converged on a synonymy-free vocabulary (with no subsequent cultural innovation in the vocabulary), the population's gene-pool can drift to reincorporate synonym-rejectors. In such a situation, the fixed cultural environment presents no problem arising from synonymy, and thus there is no selection of genotypes to cope with such problems. Here synonymy is analogous to a culturally eradicated disease; where there is no threat from the disease, genes can be reintroduced which lessen the resistance to it, with no ill effects (until the disease somehow resurfaces).

3.2 Conclusion

There may be several reasons why synonymy is rare in human languages. This model has explored the dynamics of communication involving a many-to-many mapping between a fixed set of meanings and a set of forms. The model reveals a systematic relationship between synonymy/homonymy and the transmission or reception of messages. It would be interesting to explore the possibility of a changing set of available meanings; new, more specialized, or perhaps metaphorical, meanings could be constructed out of a basic set. Such meaning-creation is certainly active in the dynamics of human languages; it is not incompatible with the present model. Economy of storage might be another factor in synonymy avoidance – why remember more than one word for the same meaning?

Computers are ideally adapted as receivers of instructions. That is why computer languages, unlike human languages, abhor homonymy. There may be various factors contributing to the rarity of synonyms in human languages. This model builds one piece of evidence that humans evolved to be well adapted as senders of messages; accurate reception of messages was less important in our prehistory. We may be primarily speakers, and secondarily listeners.

References

1. Ackley, D., Littman, M.: Altruism in the Evolution of Communication. *Artificial Life*, **4** (1994) 40–48.
2. Clark, E. V.: On the Pragmatics of Contrast. *Journal of Child Language*, **17** (1990) 417–431.
3. Hurford, J. R.: Biological Evolution of the Saussurean Sign as a Component of the Language Acquisition Device. *Lingua*, **77** (1989) 187–222.
4. Kirby, S.: Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation* **5(2)** (2001) 102–110.

5. Kirby, S., Hurford, J.R.: Learning, culture and evolution in the origin of linguistic constraints. In *Fourth European Conference on Artificial Life*, edited by P. Husbands and I. Harvey. MIT Press, Cambridge, MA. (1997)
6. Kirby, S., Hurford, J.R.: The Emergence of Linguistic Structure: An overview of the Iterated Learning Model. In *Simulating the Evolution of Language*, edited by A. Cangelosi and D. Parisi, Springer Verlag. (2002)
7. Markman, E. M.: *Categorization and Naming in Children: Problems of Induction.* MIT Press, Cambridge MA. [esp. Chs 8 & 9] (1989)
8. Oliphant, M., Batali, J.: Learning and the Emergence of Coordinated Communication. *Center for Research on Language Newsletter*, **11(1)**. (University of California at San Diego.) (1997)
9. Pinker, S.: *Language Learnability and Language Development*, Harvard University Press, Cambridge MA. (1984)
10. Smith, K.: The Importance of Rapid Cultural Convergence in the Evolution of Learned Symbolic Communication. *Advances in Artificial Life: Proceedings of the 6th European Conference on Artificial Life*, (2001) 637–640.
11. Smith, K.: The Cultural Evolution of Communication in a Population of Neural Networks. *Connection Science*, **14(1)** (2002) 65–84.
12. Smith, K.: *The Transmission of Language: models of biological and cultural evolution*, PhD thesis, University of Edinburgh. (2003)
13. Steels, L.: *The Talking Heads Experiment*, Volume 1. Words and Meanings. Antwerpen: Laboratorium. (1999) Special pre-edition.
14. Steels, L., Kaplan, F., McIntyre, A., van Looveren, J.: Crucial Factors in the Origins of Word-Meaning. In *The Transition to Language*, edited by A. Wray. Oxford University Press, Oxford. (2002)
15. Werner, G., Dyer, M.: Evolution of Communication in Artificial Organisms. *Artificial Life*, **2** (1992) 659–687.
16. Wexler, P., Culicover, P.: *Formal Principles of Language Acquisition*, MIT Press, Cambridge MA. (1980)