Intelligent Meaning Creation in a Clumpy World Helps Communication

Andrew D. M. Smith

Language Evolution and
    Computation Research Unit
Theoretical and Applied
    Linguistics
School of Philosophy,
    Psychology and Language
    Sciences
University of Edinburgh
Adam Ferguson Building
40 George Square
Edinburgh, EH8 9LL
United Kingdom
andrew@ling.ed.ac.uk

Abstract This article investigates the problem of how language learners decipher what words mean. In many recent models of language evolution, agents are provided with innate meanings a priori and explicitly transfer them to each other as part of the communication process. By contrast, I investigate how successful communication systems can emerge without innate or transferable meanings, and show that this is dependent on the agents developing highly synchronized conceptual systems. I present experiments with various cognitive, communicative, and environmental factors which affect the likelihood of agents achieving meaning synchronization and demonstrate that an intelligent meaning creation strategy in a clumpy world leads to the highest level of meaning similarity between agents.

1 Introduction

Attempts to explain the particular structure of language often appeal to a "conventional neo-Darwinian process" [21], whereby humans have evolved an innate, genetically encoded language device in the brain which is specifically tailored to the acquisition and maintenance of language [5]. More recently, however, researchers have begun to develop models which emphasize the repeated process of language learning and use it as the driving force behind the emergence of linguistic structures. For example, Kirby [12] explores in detail how certain language universals [9] can be explained elegantly by focusing on how processing complexity affects the transmission of language.

Much recent work in the field of language evolution has focused on the evolution of syntactic structure as the crucial event which marks both the genesis of language and the defining criterion which separates it from animal communication systems. Kirby [13], for example, demonstrates that syntax can arise from unstructured communication systems by creating generalized rules from the analysis of signal-meaning pairs, and Brighton [4] shows that pressures such as the poverty of the stimulus [5] lead to the emergence of syntactic structure when the process of language production and learning is repeated over generations.

There are, however, some major problems with the assumptions behind simulations such as these. Firstly, syntax develops only because signals in the simulations are coupled to pre-existing, innate, structured meanings, and so it is no surprise to find that the structure of the emergent syntax directly parallels that of the predefined semantics, as discussed by Nehaniv [19]. Explanations of the origin of these meanings, and of how they become associated with signals, are conspicuously absent. Secondly, communication consists of the simultaneous transfer of signals and meanings; thus the simulations

ignore one of the most crucial features of real language acquisition, namely that meanings are *not* transferred with words, and yet learners do manage to infer meanings and associate words with them. Thirdly, the simulations rely on variants of reinforcement learning to guide the agents [26], although the existence of reliable error signals in language learning is widely rejected [3]. In contrast, I argue that constructing meanings and learning which of them are most relevant is a crucial part of the language learning process which should not be overlooked.

The article is divided into six main parts. In Section 2, I discuss the assumption of explicit meaning transfer and its implications for models of communication and learning. In Section 3, I report details of the model of meaning creation and communication, describing how the problem of explicit meaning transfer can be overcome. In Section 4, I show the importance of meaning similarity for the emergence of successful communicative systems, and describe a baseline for meaning similarity. Finally, in Sections 5–7, I investigate how cognitive biases, communicative biases, and environmental factors such as the agents' experience and the structure of the world affect levels of meaning similarity, and therefore levels of successful communication.

2 Explicit Meaning Transfer

Kirby [13] and Batali [2] have shown separately how the simple ability to create general rules, by taking advantage of coincidental correspondences between parts of utterances and parts of meanings, can result in the emergence of a compositional, syntactic communication system. In a nutshell, this occurs when the agents are subject to pressures which limit their exposure to the language, such as the poverty of the stimulus; general rules can generate more utterances than idiosyncratic rules, are more likely to be encountered, and are therefore replicated in greater numbers in following generations. I have already noted, however, that the successful emergence of syntax in these models is dependent on the signals being coupled to structured meanings. The structure of the meanings is assumed by the model, and it is not coincidental that the syntactic structure which emerges parallels exactly the pre-existing semantic structure.

At the heart of any kind of communication system is what constitutes observable behavior during linguistic transfer, or what is actually transmitted between speakers and hearers. In Figure 1, which represents the linguistic transfer in a standard model, we can see that the speaker (on the left of the picture) utters a signal "zknvrt," but that simultaneously, the meaning in the speaker's brain (represented by three apples) is transferred directly to the hearer's brain. The hearer learns the association between signal and meaning, and crucially, it knows that *this association* is appropriate to make because the signal and meaning are explicitly linked in each communicative episode.

This kind of model of associative learning sidesteps one of the most important and difficult problems facing researchers into the acquisition of language, namely Quine's [22] famous *gavagai* problem of determining the meaning of an unfamiliar word from a set which is, in principle, infinite. The consequences of this idealization of the learning process are considerable, not least because if meanings are explicitly and accurately transferable by telepathy as in Figure 1, then the signals are not being used to convey meaning. If the signals do not convey meaning, then their role in the model is far from obvious. In fact, we can see that the inclusion of signals in the model is a complicating factor, and yet removing them brings us uncomfortably close to creating a model which bears very little resemblance to a languagelike communication system. We are left, therefore, with the conclusion that meanings *cannot* be explicitly transferred, but must instead be inferred by the hearer from the signal and the context in which they are heard.
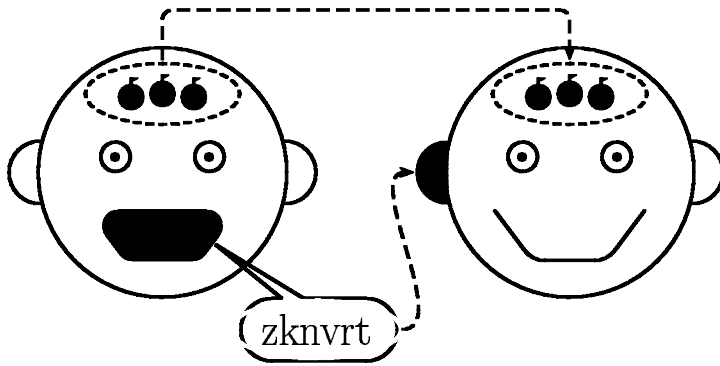
Figure 1. A communicative episode which consists of the explicit transfer of both a signal "zknvrt" and a meaning "three apples" from speaker to hearer.

So how, then, does a hearer know which meaning to associate with a signal, and where do the private meanings it uses come from? Firstly, if it is assumed that meanings are not transferable, then the agents must be able at least to infer them from elsewhere. I assume that the obvious, and most general, source for this is the world around the agent, or the environment in which it is placed. This in turn suggests that at least some of the meanings which agents talk about are used to refer to objects and events which actually happen in the environment. Binding the subjects of communication to events in the agents' world means that the agents' meanings are grounded in the world [10].

It is worth noting that the need to infer meanings from the environment has interesting implications for models such as those described by Kirby [13] and Batali [2]. These models contain no environment, and indeed nothing accessible and external to the agents, so the "meanings" used must necessarily be abstract, predefined tokens. Because they can have no reference (cannot identify any thing in the world), they cannot be inferred, and so can only be communicated through explicit transfer. In order to avoid explicit meaning transfer, therefore, there must be some kind of external world for the agents to experience in the model.

The existence of an external world in itself, however, does not mean that the problem of explicit meaning transfer is automatically avoided; for this there must be at least three separate levels of representation in the model: the external, public world, a private, agent-specific internal semantic representation, and a set of signals, which can again be publicly observed. The mappings between the public and private sections of the model must be specific to each agent and unobservable to the others; otherwise the private representations become public, making the signals unnecessary.

In Hutchins and Hazlehurst's famous neural network model of the development of a shared vocabulary [11], for instance, there is an external world made up of events, or "scenes." These scenes, however, are themselves used as the meanings for which the agents learn signals; although they are not explicitly transferred, they are publicly accessible in the communication process, and there is therefore no level of the model which is private to each agent. Brighton [4], too, presents a model with an external world made up of communicatively relevant situations. But although the environment is defined as the source of the meanings used by the agents, this relationship plays no role in the simulations; the agents never interact with the environment, and the mapping from environment to meanings is predetermined and identical for all agents. Again, there is no private level in the model, and the environment is effectively merely a complicating factor in the simulation.

Secondly, there are two possible explanations for how the agents come to have meanings which refer to things: either the meanings are innate, and have somehow evolved biologically, or they are created by the agents themselves, as a result of their interactions with the environment. Innate meanings are not inherently implausible, and they are used as a simplification in many models of aspects of language evolution (see for instance Arita and Koyama [1]), but they seem in reality to require either that the number of meanings useful to the agents be small and fixed, or that the world in which the agents exist be very stable and unchanging. If the world is dynamic, then the agents may have evolved innate meanings for something that was useful to their ancestors, but these may not be of use to them now. In practice, then, I assume that it is more reasonable to assume that the agents create meanings *de novo* in each generation, based on empirical testing of their environment, to discover which distinctions are communicatively relevant.

This paper, therefore, departs from previous accounts, which assume that language learning is equivalent to learning a mapping between signals and predefined meanings. Instead, I argue that there are at least three necessary levels of representation: a public environment, a private semantic representation, and public signals. Language learning involves the empirical creation of private meanings based on the environment, learning which of these meanings are relevant, and learning the mapping between signals and the relevant meanings which underpins communication.

3   Details of the Model

3.1   Meaning Creation

My model of independent, grounded meaning creation is based on that described by Steels [25]. I establish a simple world made up of a number of objects, which can be described in terms of the values of their *features*. In the results reported here, the world contains twenty objects unless otherwise specified. Feature values in the model are real numbers, pseudo-randomly generated in the range [0,1]. These features are abstract and do not have any specified meaning in the model, but can be profitably thought of in terms of perceptual features such as smell or color. The agents in the world interact with the objects using *sensory channels*. They have the same number of sensory channels as the objects have features, and there is a one-to-one mapping between channels and features. Sensory channels are sensitive to the feature values, and in particular can detect whether a particular feature value falls between two bounds. Meaning creation happens by splitting the sensitivity range of a channel into two discrete segments, resulting in two separate categories, or *meanings*, each sensitive to half the original range. After repeated splitting or refinement, we can represent the semantic structure on a dendrogram, as shown in Figure 2, where the nodes on the tree represent the meanings.

The agents interact with their environment through *discrimination games* [25], in which they try to distinguish one particular randomly chosen object from a context of five randomly chosen objects through the following algorithm:

- The agent investigates all its sensory channels to categorize all the objects in the context.
- If the target object is uniquely identified by any single category, then this meaning is called the *discriminatory meaning* and the game succeeds.
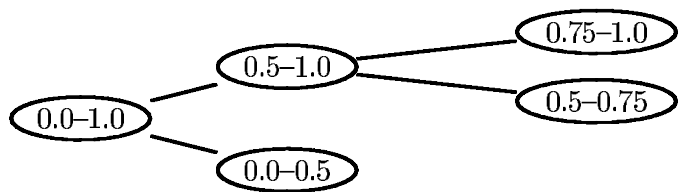- If the game fails, the agent refines a randomly chosen sensory channel.

Figure 2. Meanings represented on a Steelsian dendrogram, which has been refined twice. Each node on the tree shows the bounds between which it is sensitive.

Table 1. The categorization of objects during a discrimination game. Meanings are given in the notation c–p, where c identifies the sensory channel and p traces the path along the discrimination tree from the root to the node in question, with 0 signifying a lower branch and 1 an upper branch.

Object	Categories/Meanings		
	Channel 0	1	2
A	0–0	1–00	2–111
B	0–11	1–1	2–110
C	0–0	1–1	2–111
D	0–10	1–01	2–10
E	0–10	1–00	2–0

Table 1 shows an agent's categorization of objects during a discrimination game; the agent is investigating five objects, and has three sensory channels on which the objects are being categorized. If the aim of this game is to discriminate B from the context $ACDE$, then the game can succeed, as both 0–11 and 2–110 are possible discriminatory meanings. On the other hand, if the aim is to distinguish C from the context $ABDE$, then the game will fail, as there is no single category into which C falls which distinguishes it from all the other objects.

Failure in such a discrimination game triggers the refinement of a randomly chosen sensory channel, and therefore the creation of another level of conceptual structure in the agent. Because the sensory channel is chosen randomly, the newly created meanings may be, but are not necessarily, useful for future discrimination games. Given enough discrimination games in a static world, the agents will always develop a successful conceptual structure, although the precise details of this structure are of course not fixed, and will vary between agents and between runs of the simulation.

This semantic representation has an obvious hierarchical structure, allowing the immediate use of real semantic sense relationships such as hyponymy and antonymy to be investigated, which are not readily available in other representations. Meanings nearer the root of the tree are clearly more general than those nearer the leaves of the tree, which are more specific. Concept creation is clearly directly driven by the agents' interactions with their world, so that the meanings are not imposed from outside. The agents, therefore, have a mechanism for constructing concepts which is grounded in the environment, is based on experience, creates meanings which are useful to the agents in allowing them to discriminate between the objects they find, and results in conceptual structure which can be measured and compared. We quantify the similarity of two agents' meaning structures by averaging the similarity of the particular discrimination trees built on each of their sensory channels in turn. In greater detail, if $k(t, u)$ is the number of nodes which trees t and u have in common, and $n(t)$ is the total number of nodes on tree t, then we describe the similarity between any two trees t

and u using the following formula:

$$\tau(t, u) = \frac{1}{2}\left(\frac{k(t, u)}{n(t)} + \frac{k(t, u)}{n(u)}\right) \tag{1}$$

We can use this general measure of tree similarity τ to develop an overall measure of *meaning similarity* σ between two agents, by averaging over all their sensory channels. If $a_i t_j$ identifies channel j on agent i, and each agent has c sensory channels, then the meaning similarity σ between agents a_1 and a_2 is defined as follows:

$$\sigma(a_1, a_2) = \frac{1}{c}\sum_{i=0}^{c-1}\tau(a_1 t_i, a_2 t_i) \tag{2}$$

If two agents a_1 and a_2 have identical conceptual structures, where $\sigma(a_1, a_2) = 1$, then we refer to their meanings as being *synchronized*.

3.2 Communication

In this section, I extend the meaning creation model to investigate whether the agents can communicate with each other, using the meanings they have constructed. In order to simulate communication between the agents, I endow them with the ability to create signals, or words, which they use to express the meanings. I assume, for simplicity, that the agents can both express and understand these words without difficulty, that is, that the signals can be transmitted without error. The agents also have a dynamic lexicon of associations between words and meanings, which they use both to decide which signals to send, and to decide on an interpretation for the signals they receive. Each entry in the lexicon contains a signal s, a meaning m, a count u of how many times the pair has been used, and a confidence probability p representing the agent's confidence in the association between the signal and meaning, or the proportion of times in which s has been used that it has been associated with m. More formally, $p(s, m)$ can be expressed as

$$p(s, m) = \frac{u(s, m)}{\sum_{i=1}^{l} u(s, i)} \tag{3}$$

where l is the number of entries in the lexicon.[1]

Having successfully undertaken a discrimination game and found a discriminatory meaning, one agent (the *speaker*) utters a signal which represents this meaning. A second agent (the *hearer*) receives the signal together with the original context of objects used by the speaker. The hearer does not know which object was the speaker's target object, but tries despite this to infer the intended meaning solely from the context and from its own previous experiential history, stored in its lexicon as described above. Having inferred a meaning, the hearer then deduces the object to which it thinks the speaker was referring; successful communication occurs when the speaker's original target object is the same object as that which is identified by the hearer's meaning. It is not necessary that the agents use the same agent-internal meaning, only that both agents *refer* to the same object, or pick out the same object in the world. Importantly, neither speaker nor hearer is given *any* feedback on whether the meaning was successfully interpreted.

---

1   Further details of this communication model and of the structure of the agents' lexicons can be found in [23].

This kind of communicative model, therefore, relies neither on the explicit transfer of meaning nor on feedback to guide the learning. The algorithms for deciding which signal to choose to express a meaning, and for deciding which meaning to interpret a signal as, are therefore crucial to the success of the model. Oliphant and Batali [20] have demonstrated an ideal strategy for achieving an accurate communication system between two agents under these circumstances, which they dub *obverter*. Essentially, this strategy boils down to the speaker choosing signals which it knows the hearer will understand correctly. Unfortunately, true obverter learning assumes that the speaker has access to the lexicons of the other members of the population, so that it can choose the optimal signal for each meaning. Such mind-reading is of course unrealistic, and more damagingly returns us to a telepathic world in which communication using signals is not actually necessary. In order to avoid this, we modify the obverter strategy, by allowing the agent to read only *its own* mind, and using this as a basis for decision making; the speaker therefore chooses the signal that *it itself* would be most likely to understand if it heard the signal in this context.

The hearer, on the other hand, on hearing a signal, has only one source of information apart from the signal itself: the context in which the word was heard. It knows neither the target object to which the speaker is referring, nor the meaning which the speaker has in mind for the signal. The hearer creates a list of *possible meanings*, namely every meaning in its conceptual structure which identifies *any one* of the objects in the context and distinguishes it from all the other objects in the context. The hearer has no reason to prefer any one of these possible meanings over another yet, so each of them is paired with the signal and *lexicalized*, that is, its usage and confidence probabilities in the lexicon are updated. Once all the possible meanings have been lexicalized, the hearer searches through the list of possible meanings, and chooses the one in which it has the highest confidence. If the agent has equally high confidence in more than one meaning, then it chooses one of those meanings at random. The object which this meaning identifies is then compared with the original target object of the speaker's discrimination game, to determine the success of the communicative episode. Neither agent receives any information, however, about the success or failure of the episode.

3.3   Meaning Structure and Communication

Before investigating the interactions between meaning creation and communication, we need to verify that the modified obverter strategy can deliver successful communication without explicit meaning transfer. In order to do this, we therefore temporarily dispense with the meaning creation algorithms, and instead predefine the agents' conceptual systems. Figure 3 shows the communicative success rates for two agents whose meanings have a similarity measure of 80% (left) ($\sigma = 0.8$), and for two agents with identical, synchronized meanings (right) ($\sigma = 1$). The communicative success rate is the proportion of communicative episodes in which the target object described by the speaker is identified by the hearer.

We can immediately see on the right of Figure 3 that when $\sigma = 1$, the communicative success rate rises rapidly from zero, stabilizing as it approaches 1. In principle, the success rate will reach 1, but this is not guaranteed in a particular population over a finite time scale. On the left of Figure 3, we see that when $\sigma = 0.8$, the communicative success rate again rises rapidly in the initial period, and then stabilizes around the level of σ. Given an infinite time scale, we can expect the communicative success rate to equal the agent meaning similarity, and even over a finite time scale it forms a good approximation.

Figure 3 shows very clearly the strong link between the level of meaning similarity and the rate of successful communication. As we have eliminated both explicit meaning
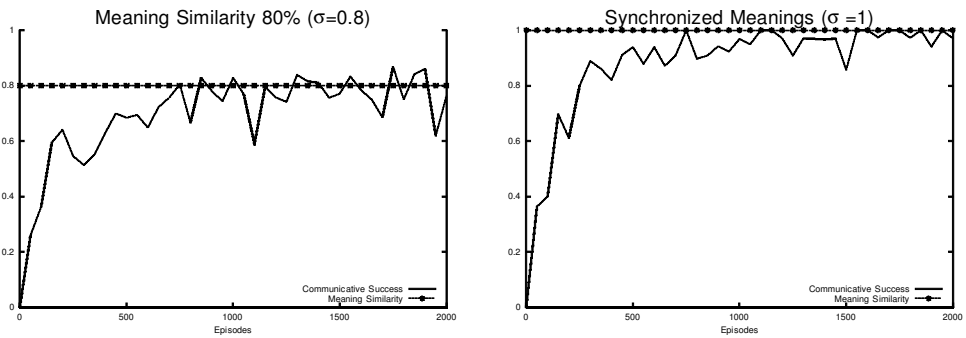
Figure 3. Levels of meaning similarity and communicative success.

transfer and also feedback from the agents to guide their interlocutors to the "correct" answer, unlike models such as those described by Steels and Kaplan [26], we force the agents to infer the meanings of words from the set of possible meanings in each context. It is clear that it is impossible for an agent to attach a word to a meaning which does not exist in its conceptual structure, and so we find inevitably that only those words which refer to shared concepts are successfully used in communication. I have also shown previously [23] how words referring to unshared meanings inevitably suffer semantic drift over time, such that they come to refer to more general meanings which are shared by the agents.

Agents, therefore, can learn communication systems without the explicit transfer of meanings, without knowledge of the topic of conversation, and without feedback about the success of the conversation guiding to the correct meaning. Successful communication arises by the context-driven disambiguation of signals, as long as agents can infer meaning from their experiences in the world. The level of communicative success is very strongly dependent on the level of meaning similarity shared by speaker and hearer.

4   The Standard (or Unbiased) Model

We have seen the importance of synchronized conceptual structure for the development of successful communication without explicit meaning transfer, but how likely is it that synchronization will occur? In this section I investigate the levels of meaning similarity, and by implication communicative success, achieved in a standard, unbiased model. This will also provide a baseline with which to compare the effects of adding cognitive and communicative biases to the agents, as well as external environmental factors such as the structure of the world and the experiences of the agents. The standard model is built on a world with two agents and twenty randomly generated objects. Each object is described in terms of ten features, and each agent has ten corresponding sensory channels on which it can build discrimination trees. The agents play a fixed number of discrimination games, with each agent having an equal probability of being chosen to play the discrimination game. There are five objects in the context, including the target object, unless otherwise stated.

If the size of the context increases, each discrimination game becomes a closer approximation to picking out one individual object from the complete set of objects in the world. An undesirable consequence of this is that the meanings created also identify particular objects in the world. In real human languages, however, words (except possibly some names) do not identify *individuals*, but rather *kinds* [8]. Experiments
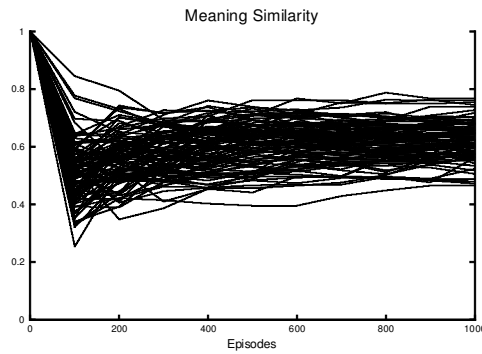
Figure 4. Agent meaning similarity (σ) rates in the standard world. 100 runs overlaid, with each run represented by one line on the graph. The mean ($\bar{\sigma}$) at 1000 episodes is 0.62 (0.61–0.64), with a coefficient of variation of 0.10.

have shown that a level around five provides a suitable balance between developing meanings which identify individuals (with large contexts) and providing the agents with too much information (with small contexts).

Figure 4 shows the level of meaning similarity between the two agents. We can see that overall there is a moderate amount of variation, with no runs producing very high or very low levels of meaning similarity. Meaning similarity is always artificially high at the beginning of each run, because both agents have sensory channels without any tree growth, and therefore identical conceptual structure. As the agents fail in the discrimination tasks, and create new meanings which are not necessarily the same as each other's, overall levels of meaning similarity fall. They then stabilize when the agents have created sufficient conceptual structure to succeed in the discrimination tasks, and there is no further need for much meaning creation. To measure the relative variation we see in Figure 4, I have taken a cutoff point of 1000 episodes, and calculated the average (mean) agent meaning similarity $\bar{\sigma}$ and the coefficient of variation (CoV), which is the standard deviation expressed as a percentage of the mean.[2] I express $\bar{\sigma}$ together with a 95% confidence interval, recognizing that the particular 100 runs of the simulation we have carried out only represent a sample drawn from an infinite set of runs. In the standard model, therefore, we expect to get meaning similarity rates of about 62%, which is not high enough to produce a very successful communication system under normal circumstances. In the following sections, I investigate how variations on this standard model will affect the levels of meaning similarity which the agents achieve.

5   Cognitive Biases and Tree Growth Strategies

In order to explain the apparent paradox of child language acquisition, researchers have regularly appealed to several particular cognitive biases, including the *object bias* [16], which states that a child will assume that an unfamiliar word names a whole object, rather than a particular property of it, and the *shape bias* [14], which states that a child is more likely to assume that an unfamiliar word refers to the shape of an object rather than to other properties such as its color or taste. In our model, the channels are intrinsically meaningless, so we cannot speak in terms of particular properties, but we can investigate how more abstract biases affect the construction of conceptual categories.

---

2 The standard deviation is scaled relative to the mean so that we can more accurately compare results from distributions with different means.

When a discrimination game fails, the agent chooses a channel on which a node will be refined. This is done on the basis of the channel's bias b_{a_n}, where a identifies the agent and n the number of that agent's sensory channel. The bias is specified when the agent is "born," and does not change during the simulation; it is equivalent to the probability of channel n being chosen for refinement. In the standard model, each channel bias is the same (i.e., there is a uniform bias distribution), and so the agent essentially chooses a channel at random each time, but the channel biases can of course be defined according to particular probability distributions. We will now look at *random* biases, where the bias for each channel is chosen randomly at the start of the simulation; and *proportional* biases, which are defined according to a fixed probability distribution. With proportional bias allocation, the bias on each channel represents a fixed proportion p of the remaining bias, taking into account biases which have already been allocated, as follows:

$$\text{if} \quad n = 0, \qquad b_{a_n} = p$$
$$\text{if} \quad n > 0, \qquad b_{a_n} = p \left(1 - \sum_{i=a_0}^{a_{n-1}} b_i \right) \tag{4}$$

Because the biases represent probabilities, they are always scaled after allocation so that the sum of biases for each agent equals 1. For instance, if p were 0.5, and the agent had five channels, then the biases would be allocated as in Table 2. We can also see that the allocation of biases by proportions is deterministic, so if two agents have the same value of p, then they will have identical cognitive biases. Unless specified otherwise, p is set to 0.5 for all simulations reported here. Under proportional bias allocation, channels with lower numbers always have higher biases, but this is purely an artefact of the implementation, and nothing in the results relies on it.

As well as changing the biases, and therefore the likelihood of tree growth occurring on particular channels, we can also define completely different strategies for the channel choice. In addition to the *probabilistic* method, where the agent chooses a channel at random based on the biases described above, we will investigate another strategy, when the agent searches through its channels in order of their biases, until it finds a refinement which would have resulted in successful discrimination *in this particular discrimination game*, had the refinement already taken place. If no channel which meets this criterion is found, then no refinement takes place.

A crucial feature of this strategy, which I call the *intelligent tree growth strategy*, is that a refinement will always make a helpful distinction in at least the particular discrimination game during which it was created, whereas refinements under the probabilistic strategy are not guaranteed to be successful at all.

Table 3 shows the average rate of agent meaning similarity after 1000 episodes, averaged over 100 runs of the simulations as above, with both the tree growth strategies (probabilistic and intelligent) and the channel bias allocations (uniform, random, and proportional) being varied.[3] Counterintuitively, we find that the best results are achieved under the uniform, standard model which we looked at in Figure 4. The same level is achieved if agents have proportionally allocated biases, suggesting that the important factor is that in both these cases the agents' biases are *identical*. When the agents have random biases, on the other hand, then the level of meaning similarity drops to just over 50%. Under the intelligent strategy, it is interesting that the level of

---

3 The combination of uniform biases and intelligent tree growth strategy is not included, because the intelligent tree growth strategy is based on searching the channels in order of their probabilities; if these are all equal, then there is no obvious way to order them except randomly, which makes the search equivalent to a random, or probabilistic, choice.

Table 2. Allocation of biases under the fixed proportional method, with $p = 0.5$.

Channel n	Bias b_{a_n}	Scaled Bias
0	0.5	0.5161
1	0.25	0.2581
2	0.125	0.129
3	0.0625	0.0645
4	0.03125	0.0323

Table 3. How different tree growth strategies and cognitive biases affect average agent meaning similarity rates.

Strategy	Biases	$\bar{\sigma}$	CoV
Probabilistic	Uniform	0.62	0.10
	Random	0.52	0.18
	Proportional	0.62	0.18
Intelligent	Random	0.39	0.35
	Proportional	0.43	0.30

meaning similarity is even lower, and the variation very high, with some runs producing meaning structures with almost no similarity at all.

So why do agents produce very divergent conceptual structures when they use the intelligent tree growth strategy? The intelligent strategy always focuses refinements on channels which would have succeeded, and, other things being equal, channels which already have high levels of tree growth are more likely to produce a discriminatory meaning than those which have only very general meanings. Therefore, after a few initial refinements have been made, the intelligent strategy tends to focus further refinements on those channels on which trees have already been grown, and so divergence is therefore almost inevitable under this strategy, unless the initial refinements made by the agents happen to be the same.

6   The Principle of Contrast

Biases which may help explain language acquisition are not just proposed in relation to meaning creation, but also to communication; Clark [6], for instance, proposed the *principle of contrast* (PoC), that every difference in a signal corresponds to some difference in meaning, whereas Markman [17] put forward the closely related *mutual exclusivity assumption* (MEA), that children assume that objects do not belong to more than one category. For example, Markman and Wachtel [18] describe how experimenters present children with a banana and a whisk, and then ask them to "show me the fendle." The children tend to interpret *fendle* as referring to the whisk, and it is hypothesized that this is because they already know a word for the banana, so they assume that the unfamiliar word must refer to the unfamiliar object. More recently, these suggestions have been complemented by further research showing how language itself appears, to a certain extent, to shape the learner's meaning structure despite innate biases [15].

The crucial idea underlying both the PoC and the MEA, which can be expressed simply as "every difference in a signal corresponds to some difference in meaning" and implies that there are therefore no true synonyms, can be implemented in our model by ensuring that when an unfamiliar signal is encountered, an agent will create a new meaning which corresponds to one of the objects in the context, and assume that the new signal corresponds to this meaning. This means that meaning creation
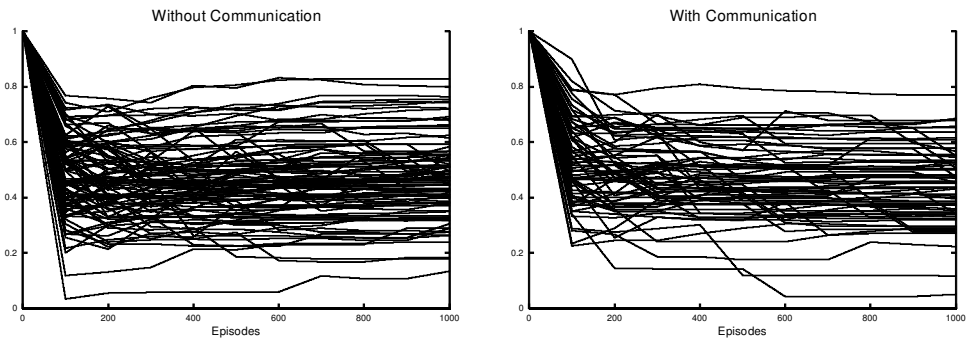
Figure 5. Meaning similarity rates with discrimination-driven meaning creation (left), $\bar{\sigma} = 0.43$ (0.41–0.46), and with the addition of communication-driven meaning creation (right), $\bar{\sigma} = 0.47$ (0.46–0.48).

can therefore now be triggered by two mechanisms in the model: not only failure in the discrimination game, but also failure in the interpretation of an unfamiliar word.

Figure 5 shows how adding meaning creation driven by failure in communication to the model actually has very little effect on the overall level of meaning similarity. We can see that there is a slight increase in $\bar{\sigma}$, but if we use the Kolmogorov-Smirnov (KS) statistic, which expresses how different two distributions are [7], we find that there is no statistical difference between the two sets of results. This would initially appear somewhat surprising, given the frequency with which such heuristics are apparently invoked in the learning of words by children, but in this model it is explained by the fact that the extra information received by the hearer when it receives an unfamiliar word, which it uses to create a new meaning, does not sufficiently help the hearer to build a conceptual structure closer to that of the other agent. Because the words created in the model do not identify individual objects, the occurrence of a new, unfamiliar word is relatively rare. Even when this does occur, the meaning creation process itself is of course unguided, so there is no guarantee that the hearer will build appropriate new conceptual structure, as there is no external pressure to maximize meaning similarity.

7 Environmental Factors

7.1 Experience
This model of empirical meaning creation is based on the agents' building their conceptual structure in response to failures in their interactions with the world, and it would seem reasonable therefore to investigate the importance of the particular situations which they experience. Humans who have similar experiences create distinctions based on those experiences which can be unnoticed or irrelevant to others who have not had them, leading to the creation of particular specialized terminology or jargon to name these distinctions.

In order to investigate how much of the agents' conceptual structure is influenced by the order in which they encounter certain objects and sets of objects, I have implemented simulations in which both agents are given identical discrimination games to perform. Each discrimination game itself still consists of a random target object to be distinguished from a random set of objects, but both agents now undertake the same discrimination game, creating meanings when they fail as in previous experiments. Table 4 shows the levels of meaning similarity achieved when the agents are given identical discrimination games to perform, compared to the results in our reference table (Table 3) when they have different, randomly chosen games. Large values of the

Table 4. How the agents' experiences affect average agent meaning similarity rates.

Strategy	Biases	$\bar{\sigma}$	
		Diff. exp.	Same exp.
Probabilistic	Uniform	0.62	0.63
	Random	0.52	0.54
	Proportional	0.62	0.64
Intelligent	Random	0.39	0.54*
	Proportional	0.43	1.00**

KS statistic show that the meaning similarity distributions are statistically significantly different; in this article, distributions where $p < 0.05$ are denoted by an asterisk (*), and those where $p < 0.01$ are denoted by a double asterisk (**).

We can clearly see that under the probabilistic strategy, there are no significant differences when the agents have identical experiences, but that in contrast, the intelligent strategy produces significantly increased levels in meaning similarity, under both random and proportional biases. Indeed, if the agents have the same biases and the same experience, we have in effect a deterministic situation, and so it is no surprise that we find complete meaning synchronization ($\sigma = 1$) in this case.

7.2 A Clumpy World

The world in which we live is not uniformly random; indeed, there are many constant properties behind the phenomena we encounter, which can be described in terms of physical and chemical laws. We know, for instance, that unsupported objects will always fall until they reach a lower surface. Scientists can measure the gravitational field which causes this, and we know that its magnitude decreases as the object moves further from the center of the planet; yet in practical terms, the objects in our world do not differ in the gravitational field applying to them. In terms of a space of possible worlds, all the objects in our world are *clumped* together in one section of the space, where the gravitational field is always constant.

Bloom [3] describes how babies use the structure in the world, such as the properties of objects, to make sense of it through categorization and, ultimately, in deciphering the meaning of words. K. Smith [24] has shown how compositional systems are more likely to emerge in generalizing agents when the environment exhibits a high degree of structure. In this model, I investigate how the agents fare in the meaning construction task in a world which is structured or constrained in certain ways, and I explore how the meaning similarity which emerges differs from that in a random world.

In a *clumpy* world, the objects are grouped together in some way and this is implemented in our model by giving each member of a group *identical* feature values for some particular feature (such as the gravitational field applying to them). This means that the objects in a particular group are therefore a priori indistinguishable on this channel, no matter how many times the discrimination tree is refined, and so the objects can only be told apart using meanings created on another sensory channel. In the random world, we could consider each object as a group in itself, with each group containing just one object; in the clumpy world, we choose the number of groups arbitrarily according to the channel and the number of objects in the world. The number of groups on channel c, $g(c)$, is taken as follows:

$$g(c) = \frac{O}{c + 1} \tag{5}$$

Table 5. Allocation of groups in a clumpy world.

Channel c	0	1	2	3	4	5	6	7	8	9
Groups $g(c)$	20	10	7	5	4	4	3	3	3	2

Table 6. How the structure of the world affects average agent meaning similarity rates.

		$\bar{\sigma}$	
Strategy	Biases	Random world	Clumpy world
Probabilistic	Uniform	0.62	0.70*
	Random	0.52	0.59*
	Proportional	0.62	0.68*
Intelligent	Random	0.39	0.82**
	Proportional	0.43	0.88**

where O is the number of objects in the world. If there is no exact division, then $g(c)$ is always rounded up to the next whole number.

In a world of 20 objects, therefore, the number of groups on each channel will be as shown in Table 5. We can see that the channels toward the end of the list have few groups, and so are much less likely to be of any use in a discrimination game, though we also note that none is completely useless if all objects fall into one group (this would only happen under this setup if the agents had more sensory channels than there were objects in the world). The groups are arbitrarily biased so that more distinctions can be made on low-numbered sensory channels, just as the proportional allocation of biases was biased toward low-numbered sensory channels. If the structure of the world is biased in a certain direction, it makes sense, if we want to appeal to some selectionist motivation for the existence of the cognitive biases, for the channels to be biased in a similar way.

Table 6 shows that all tree growth strategies produce significantly higher levels of meaning similarity than in simulations under the same conditions in a uniformly random world. The probabilistic strategy produces significantly increased levels of meaning similarity under all conditions where the order of the agents' experiences did not have any significant effect. Under the intelligent strategy, the levels of meaning similarity have more than doubled in comparison with those achieved in the uniformly random world, and the differences are highly statistically significant ($p < 0.01$).

An intelligent meaning creation strategy, therefore, results in poor meaning similarity levels if the agents are in a random world, but it is very good at taking advantage of any structure in the world, and produces very high meaning similarity levels in a clumpy world.

8 Summary

In this article, I have described a model of empirical meaning creation and of the evolution of communication, in which successful communication can emerge without innate meanings and without the explicit transfer of meanings; I have also described the importance of meaning synchronization in the model. Furthermore, I have investigated meaning similarity levels under various conditions, experimenting with various cognitive, communicative, and environmental factors, motivated by research into how children acquire and learn what words mean.

The structure of the world plays a large role in determining which strategy of meaning creation will create a conceptual structure which is most likely to result in successful

communication. If the objects in the world are distributed randomly, then the agents can do no better than create meanings based on their innate biases, and reasonably high similarity will occur when the agents happen to have the same biases. If the world is structured, on the other hand, then it is much better for the agents to use an intelligent strategy for meaning creation, which takes account of the structure in the world to a much greater degree.

References

1. Arita, T., & Koyama, Y. (1998). Evolution of linguistic diversity in a simple communication system. *Artificial Life, 4,* 109–124.

2. Batali, J. (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In E. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models* (pp. 111–172). Cambridge, UK: Cambridge University Press.

3. Bloom, P. (2000). *How children learn the meanings of words.* Cambridge, MA: MIT Press.

4. Brighton, H. (2002). Compositional syntax from cultural transmission. *Artificial Life, 8,* 25–54.

5. Chomsky, N. (1980). *Rules and representations.* London: Basil Blackwell.

6. Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Erlbaum.

7. Conover, W. J. (1999). *Practical nonparametric statistics*, 3rd ed. New York: Wiley.

8. Cruse, D. A. (1986). *Lexical semantics.* Cambridge, UK: Cambridge University Press.

9. Greenberg, J. H. (1966). *Universals of language*, 2nd ed. Cambridge, MA: MIT Press.

10. Harnad, S. (1990). The symbol grounding problem. *Physica D, 42,* 335–346.

11. Hutchins, E., & Hazlehurst, B. (1995). How to invent a lexicon: The development of shared symbols in interaction. In N. Gilbert & R. Conte (Eds.), *Artificial societies: The computer simulation of social life* (pp. 157–189). London: UCL Press.

12. Kirby, S. (1999). *Function, selection and innateness: The emergence of language universals.* Oxford, UK: Oxford University Press.

13. Kirby, S. (2002). Learning, bottlenecks and the evolution of recursive syntax. In E. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models* (pp. 173–203). Cambridge, UK: Cambridge University Press.

14. Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3,* 299–321.

15. Levinson, S. C. (2001). Covariation between spatial language and cognition, and its implications for language learning. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 566–588). Cambridge, UK: Cambridge University Press.

16. Macnamara, J. (1972). The cognitive basis of language learning in infants. *Psychological Review, 79,* 1–13.

17. Markman, E. M. (1989). *Categorization and naming in children: Problems of induction. Learning, development and conceptual change.* Cambridge, MA: MIT Press.

18. Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology, 20,* 121–157.

19. Nehaniv, C. L. (2000). The making of meaning in societies: Semiotic and information-theoretic background to the evolution of communication. In *Proceedings of the AISB Symposium: Starting from Society—The application of social analogies to computational systems, 19–20 April 2000* (pp. 73–84). AISB.

20. Oliphant, M., & Batali, J. (1997). Learning and the emergence of coordinated communication. *Center for Research on Language Newsletter*, *11*.

21. Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, *13*, 707–784.

22. Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

23. Smith, A. D. M. (2001). Establishing communication systems without explicit meaning transmission. In J. Kelemen & P. Sosík (Eds.), *Advances in artificial life: Proceedings of the 6th European Conference on Artificial Life*, Lecture Notes in Artificial Intelligence, 2159 (pp. 381–390). Heidelberg: Springer-Verlag.

24. Smith, K. (2002). *Compositionality from culture: The role of environment structure and learning bias* (Technical report). Theoretical and Applied Linguistics, University of Edinburgh.

25. Steels, L. (1996). Perceptually grounded meaning creation. In M. Tokoro (Ed.), *Proceedings of the Second International Conference on MultiAgent Systems, ICMAS-96* (pp. 338–344). Menlo Park, CA: AAAI Press.

26. Steels, L., & Kaplan, F. (2002). Bootstrapping grounded word semantics. In E. Briscoe (Ed.), *Linguistic evolution through language acquisition: Formal and computational models* (pp. 53–73). Cambridge, UK: Cambridge University Press.