On the Origin of Negation[*]

Giorgio Sbardolini

ILLC and Philosophy Department, University of Amsterdam

November 18, 2022

Abstract

The ability to express negation in language may have been the result of an adaptive process. However, there are different accounts of adaptation in linguistics, and more than one of them may describe the case of negation. In this paper, I distinguish different versions of the claim that negation is adaptive and defend a proposal, based on recent work by Steinert-Threlkeld (2016) and Incurvati and Sbardolini (2021), on which negation is an indirect adaptation.

1 Introduction

According to a popular metaphor, languages are biological organisms: complex systems composed of parts that hang together in a functional whole. The metaphor mantains that words stand to languages as organs stand to organisms: components that can be identified by their primary function and that combine with other components to ensure that the whole system operates as it should: by supporting life in one case, and communication in the other.[1]

Pushing the metaphor further, perhaps the laws of change that govern evolution in biology apply to language as well. Like living organisms, languages change over time, with new parts being developed and other parts becoming irrelevant and eventually disappearing. Whole new languages sometimes emerge, while others may go extinct. Clearly, there are disanalogies between biological and linguistic change: for one, the time scales are incomparable, as Latin is now gone but was spoken by members of our species. Indeed, the whole metaphor of languages as organisms may not mean much. Nevertheless, some distinctions that are important for our understanding of biological evolution are also relevant for thinking about the evolution of words and of the concepts they express. An example is the distinction between direct and indirect adaptation.[2] This paper has two goals. The first is to stress the importance of this distinction for linguistic evolution, and in particular

---

[1]Richard (2019) appears to be an extended argument in favor of the language/organism metaphor. Darwin (1871) himself found the analogy between linguistic and biological evolution somewhat tempting (p. 60). Aronoff (2017) includes a brief history of the analogy and of its remarkable impact on the development of linguistics. See also Richards (2002).

[2]On adaptationism in biology, see Godfrey-Smith (2001).

for the evolution of negation. The second is to argue that Incurvati and Sbardolini's (2021) recent account of the evolution of negation, complemented with the account of compositionality in Steinert-Threlkeld (2016), offers a plausible model of negation as an indirect adaptation.

I will outline the distinction between direct and indirect adaptation in §2. This will take us to discuss the relationship between biological and linguistic evolution in §3. There are different versions of the claim that the evolution of negation in human language is an adaptive process. I will discuss two views on which negation is a direct adaptation, in §4 and in §5, and find neither convincing. In §6 I will review the account of Incurvati and Sbardolini (2021), on which, I argue, negation is an indirect adaptation. This third account avoids the difficulties of the previous two, and stands to illustrate the importance of the distinction between direct and indirect adaptive processes in linguistics. In combination with a recent compositional account of the evolution of negation (Steinert-Threlkeld, 2016), I conclude with the outline of a plausible story about the origin of negation.

2   Direct and Indirect Adaptation

From an evolutionary perspective, wings are quite mysterious. Wings are used to fly, and they seem marvelous adaptations for a life in the skies. Moreover, wings must be explainable in evolutionary terms, for we see them develop independently in unrelated species: in insects and birds, but also in reptiles (the pterosaurus) and mammals (bats). As Darwinian evolution has it, the organs and parts of an animal developed over millions of years through imperceptibly small modifications till their current form, structure, and function. How could wings ever develop? Certainly wingless animals could not have produced a winged offspring in the space of a generation, or we should expect a Pegasus born of terrestrial horses. But what should be the use of a proto-wing too small for takeoff and landing?

The puzzle is that an animal must be spending significant time mid-air in order for an environmental need for flying to be of any advantage, and yet minuscule and underdeveloped wings appear to be of no marginal benefit for ground-dwelling animals. Perhaps proto-wings could be useful for gliding, or for softening a fall—but these initial hypotheses have now been refuted. What we now know is that wings used to be (and in some cases still are) devices for bodily heat control. Their development is supported by the advantages of increasing surface area with the least increase in volume. Later on, animals with large flat appendages on their backs exploited them to move around in the air (Gould, 1991).

The case of wings illustrates an important distinction between primary function (flying) and evolutionary origin (thermoregulation). The distinction underlies a difference between two processes: direct and indirect adaptation. A feature such as an animal's organ (but also a type of behavior, or a physical property that does not anatomically count as an organ) is the result of direct adaptation if its development has been supported by its current primary use. This may be the case of the eye, developed for the rapid detection of changes in light intensity. A feature is the result of indirect adaptation if its development is explained by some function other than the primary one it presently performs. This is the case of the wing. Indirectly adapted features may be solutions to challenges to the organism's survival, but they did not evolve for the function they appear to perform at later stages of the organism's development.

Moreover, some features appear by chance as mutations (and at some level, randomness pervades all biological evolution): function explains nothing of their first appearance. Nevertheless, mutations may then be co-opted for some purpose, and in this case they may persist and be passed on.

Co-option is not direct adaptation, and indeed, sometimes *co-option* is used to mean indirect adaptation. To fix terminology, in this paper co-option and indirect adaptation are distinguished, to emphasize the role that function plays in the latter process but not in the former. Finally, *exaptation* designates the result of any evolutionary process other than direct adaptation, following the terminology of Gould and Vbra (1982).[3]

Different types of evolutionary processes are observable in language evolution too. Some words are plausibly described as direct adaptations, for language is shaped by the speakers' communicative purposes: words like *water, fruit, kayak*, and so on. Moreover, chance matters—for instance when errors are passed on to the next generation of language users. This point is not fully appreciated, although philosophers are familiar with specific cases. Gareth Evans (1973) famously tells the story of *Madagascar*, which originally referred to a region of the African mainland, and now refers to the large island nation. Evans's story contains some unnecessary complications, such as the fact that the shift occurred as the word was borrowed by a new linguistic community. At the very least, the story seems to support the view that *Madagascar* is a case of co-option—analogizing the Europeans' mistaken reference to random variation.[4] What's not fully appreciated is that the distinction between types of evolutionary processes in linguistics goes well beyond the theory of names.

Another interesting example is *condescension*. In an entry written at the end of the 19th century and not revised since (Hanks, 2013, 161), the OED defines condescension as follows:

> Voluntary abnegation for the nonce of the privileges of a superior; affability to one's inferiors, with courteous disregard of difference of rank or position.

Condescending behavior used to be positively regarded. But when Victorian reformers became opposed to condescending models of government, they passed on to us the modern notion of a paternalistic and patronizing attitude (Siegel, 2005). Here the semantic shift is not due entirely to chance, although it seems clear that the word currently has a semantic function different from the one it initially had. The history of *condescension* is that of an indirect adaptation.

3   Evolution outside Biology

There is a large debate on how non-biological evolution is best understood (Mesoudi, 2011). I will introduce my discussion of negation by addressing two important questions about evolution: first, what are the laws of change? Second, what are its causes?

A biological organism is a complex whole of features that work together to ensure the organism's survival and reproductive success. The whole hangs together in a balance of costs and benefits. The benefits are tied to whatever advantages a certain feature confers to the animal. These advantages can be approximately quantified as the average number of offspring reaching sexual maturity. There are also costs, however. Some costs are due to energy consumption: the food and resources necessary for the development and maintenance of a new feature. There may be other costs: the peacock's long and colorful tail is great to attract the female but can be a serious hindrance when running

---

[3]Another term that may be familiar, especially from the popular literature, is *spandrel* (Gould and Lewontin, 1979). As I understand it, spandrels are exapted features which, in addition, need not perform a specific function.

[4]There are many familiar examples of co-option in morphology. For example, the article *al-* has mistakenly been taken to be the first syllable of many words imported in European languages from Arabic, e.g., *algebra*. The role of errors in language change is central to standard formal analyses of it—see e.g., Nowak and Krakauer (1999). On exaptation and language change, see Traugott (2004).

away from predators (Zahavi, 1975; Lachmann et al., 2001). On a competing-pressures account of change, organs developed as direct adaptations can be explained as innovations whose benefits outweigh the costs, and evolution itself as an optimization process perturbed by drift.

The general shape of this account may apply to language as well. As a natural development, the word-for-word composition of language is the result of a trade-off between information transmission (benefit) and cognitive load (cost), under the noise generated by error and external contingencies (see Kemp et al., 2018, for an overview). Minimization of cost favors languages with as few words as possible: the easiest language to memorize and acquire consists of one word and one syllable. The downside is a lot of ambiguity: communication in such a hypothetical language would be extremely inefficient. On the other hand, maximization of benefit tends to eliminate ambiguity. A hypothetical language without any ambiguity would have a different word for any subtly different shade of meaning, and information could be transmitted with the most extreme precision. On the downside, there would be so many words to commit to memory to make such a language impossible to learn and use. On these assumptions, a direct linguistic adaptation is a linguistic feature (a word, but also a morpheme or a grammatical construction) whose emergence is explained as a matter of communicative benefits outweighing cognitive costs.[5]

Languages in the natural world strike a balance between these forces, finding a compromise between the two extremes (Greenberg, 1963; Croft, 2000; Haspelmath, 2021).[6] Consider for example the lexical category of sentential coordinators. The evolutionary account of the set of coordinators in language parallels the cost-benefit analysis that may be applied to an animal's organs. English contains conjunction *and*, disjunction *or*, and negated disjunction *nor*. No language has more operators of this kind, or different ones. No language, for example, contains a simple word to express negated conjunction (*nand) or the material bi-conditional (Horn, 1972; Uegaki, 2022). All languages have at least one of conjunction and disjunction, if not both. The remaining truth-conditional relations are expressed either by combining the lexicalized coordinators compositionally, or by letting context disambiguate at the pragmatic level. Efficient communication supported the development of enough coordinators that the necessary semantic distinctions among Boolean operators could be made efficiently, but not so many that cognitive costs would explode.[7]

What are the underlying causes of optimization of function in language? On causality, as noted by Saussure (Aronoff, 2017), the analogy between linguistic and biological evolution breaks down. (Faced with this hurdle, Saussure argued that the science of language ought not to deal with evolutionary questions.) The great Darwinian insight, later confirmed by Mandelian genetics, is

---

[5]There are alternative views of language change in the philosophical literature. A prominent figure is Donald Davidson, who has emphasized the creative role of individual speakers in generating change (Davidson, 1984, 1986, 1991). According to Davidson, speakers' free individual will and private intentions sustain change (see Armstrong, 2016). A voluntaristic and individual-driven conception seems also to underlie many "engineering" accounts of change (Cappelen, 2018). These views may describe actual or possible developments, but which can hardly be understood as natural or spontaneous. Here I am focusing on the latter kind of developments.

[6]On the nature of the forces shaping language change, Zipf (1949) contrasts speaker's economy (minimize effort) and hearer's economy (maximize information), while for Martinet (1964) communicative needs conflict with articulatory and mental inertia. Setting aside a more precise characterization, these account describe language evolution as an optimization process.

[7]Some languages, like Warlpiri (Bowler, 2015) and Maricopa (Gil, 1991), contain lexical items expressing disjunction, but none for conjunction. Others, like ASL (Davidson, 2013) and Japanese (Sauerland et al., 2015) contain coordinating expressions that are ambiguous between *and* and *or*. Many languages lack disjunction, such as Wari' (Mauri, 2008). For recent accounts of this curious distribution, all along the general lines of competing costs and benefits, see Uegaki (2022); Enguehard and Spector (2021); Carcassi and Sbardolini (2022); Bar-Lev and Katzir (2022). Other work applies similar ideas to kinship terms (Kemp and Regier, 2012), color terms (Regier et al., 2007), quantifiers (Steinert-Threlkeld, 2020), person systems (Zaslavsky et al., 2021), and indefinites (Denić et al., 2022).

that sexual selection is the engine of biological change. The topic of this paper is why, possibly, specific features of human language developed, such as the presence of negation. Similar questions arise about the connectives, about words such as *water* or *kayak*, about compositionality, and so on. These cannot plausibly be all a matter of reproductive fitness. An alternative hypothesis is that they are a matter of learning. A large body of literature views language evolution as a form of cultural evolution, on which iterated learning of behavior is the driving factor of change.[8] Cultural transmission goes at least some way toward a solution to Saussure's problem of finding a non-biological basis for language change. On this account, language evolution can be described as a noisy optimization process, whose main underlying cause is cultural transmission. Within this broad framework, we can ask about the evolution of specific features of language, such as negation.

4   Facts-first adaptationism

Like the coordinators *and, or, nor*, negation *not* also expresses a truth-conditional relation, and in some form or other negation is a feature of all world languages, as a lexical and morphological category (Horn, 1989; Dryer, 2005; Miestamo, 2005).[9] Like the wings, negation seems to be a significant evolutionary "jump", but evolution is a local process. Perhaps the universal distribution of negation is evidence of adaptation. I begin by discussing two versions of the direct adaptation hypothesis: a 'facts first' and a 'frequencies first' version. Neither is convincing.

There is a general reason why adaptationism in linguistics is often found to be compelling. We have to infer causes from effects. If the effects are direct adaptations we can explain the emergence of a word as a response to a specific informational need: to convey information of a particular sort, or to talk about particular sorts of facts. According to facts-first adaptationism, we have words that can be used to efficiently exchange information of a specific sort fairly reliably. Negation may be no exception.

Some philosophers have committed to general arguments for adaptationism. For example, Daniel Dennett (1995) holds that all aspects of biological evolution are explained along adaptationist lines.

> [Adaptationism] plays a crucial role in the analysis of every biological event at every scale from the creation of the first self-replicating macromolecule on up. ... Adaptationist reasoning is not optional; it is the heart and soul of evolutionary biology. (Dennett, 1995, 238)

Dennett also claims that all evolutionary processes, including linguistic ones, proceed by the same mechanism, which operates on genes inside the cell, and on 'memes' outside of it. (Memes are supposed to be the non-biological counterparts of cells.)

> Not only all your children and your children's children, but your brainchildren and your brainchildren's brainchildren must grow from the common stock of ... genes and memes, that have so far been accumulated and conserved by the inexorable lifting algorithms ... of natural selection and its products. If this is right, then all the achievements of

---

[8]See Kirby (1999); Nowak and Krakauer (1999), and Kirby et al. (2014) for a more recent survey. An important paper for the study of iterated learning is Griffiths and Kalish (2007).

[9]Some might think that the question why negation evolved in natural language is quite trivial, since it is the expression of some kind of internal mental function of the Language of Thought (Fodor, 1975). In this case, the question is not so much how negation emerged in English or Urdu, but how it emerged in the Language of Thought. The Language of Thought hypothesis may or may not be true, but it doesn't add anything important at the moment, as the question about evolution remains.

> human culture—language, art, religion, ethics, science itself—are themselves artifacts
> (of artifacts of artifacts ...) of the same fundamental process that developed the bacteria,
> the mammals, and *Homo sapiens*. (Dennett, 1995, 144)

Dennett appears to suggest that every biological feature of an organism is adaptive, and that there is but one kind of evolutionary process common to biology and linguistics. A philosopher in agreement with Dennett would then have to explain the evolution of negation as an adaptation.[10] Does this mean that all evolutionary processes are the result of direct adaptations? This would make Dennett's claim very controversial and hard to defend. Or is this a looser use of the term, such that what I'm counting as indirect adaptations, or even cases of co-option, count as adaptations? This ambiguity leaves Dennett's general claim open to interpretation.

There could be a range of adaptationist explanations all looking very different from one another, but there is one type of account of linguistic adaptation that many people have found compelling. The account is a version of the competing-pressures model of change. It goes roughly as follows. People may be in different information states: a state of *Water is around here*, a state of *Fruit is over there*, and so on. Linguistic expressions develop over time supported by the benefits of accurate information sharing: communication is a cooperative game of speaker and listener, won by both if both come to believe that water is around here just in case water is around here, that fruit is over there just in case fruit is over there, and so on. In this game, the listener does not have direct evidence about water or fruit, and has to rely on what the speaker says, who can see which information state they're in and communicate to the listener accordingly. This is the signalling game, and the basis of David Lewis's (1969) influential account of linguistic conventions—later developed by Skyrms (2010) and others into an evolutionary account. On this picture, conventions about the use of *water* and *fruit* develop as functional responses to environmental information. More generally, words evolve as providers of an informational need to communicate about certain facts. Information about water or fruit is no doubt useful, and presumably the cognitive costs of memorizing these words are greatly compensated by the benefits of sharing the relevant information.

This sketch of the Lewis/Skyrms view is oversimplified and idealized in a number of ways. Even so, it is at least the beginning of an account of the origins of linguistic conventions for creatures like us. Many have found the Lewisian picture independently plausible. Moreover, the overall account is directly adaptationist: for the function of *water* and *fruit* is primarily to refer to water and fruit, and the direct adaptation thesis is that it is this function that carried the evolution of the words.[11] Perhaps a similar story applies to negation too.

Just like people may happen to be in a state of *Water is around here* or a state of *Fruit is over there*, so there may be *Water is not here* and *Fruit is not there* information states. Negative information must also be useful, if its positive counterpart is. Negation is ordinarily used to communicate that something is not the case, absent, or non-existent. Since evolutionary development is driven by communicative function, the word *not* (and its counterparts in Spanish, Mandarin, and

---

[10]As a reviewer points out, Dennett seems to have in mind (in the quoted passage) language as a whole, and it's unclear if he would commit to specific aspects of language, such as negation, being adaptive. For sure, adaptationism about linguistic abilities as a whole does not entail adaptationism about specific aspects of language. Nevertheless, Dennett is probably to be read as supporting adaptationism "all the way down", considering that he regards it as 'not optional' (cf. the first citation) and that he speaks of 'the same fundamental process' (second citation).

[11]I am skipping several details in order not to obscure the essential structure of the argument for direct adaptation. For example, I speak of the evolution of words (such as *water*, *fruit*, and *not*) although presumably the main linguistic unit for communication is the sentence (*Water is around here*, etc.). Further details on the Lewisian picture of conventions can be found in the work of Brian Skyrms (2010). Versions of the account are popular in cognitive science (Tomasello, 2003) and linguistics (Croft, 2000).

so on) evolved for the benefits of communicating negative facts, such as the fact that there is no water on that hill, that this fruit is not ripe, or that unicorns do not exist. The little extra cognitive effort needed to develop negation must have proved to be no impediment.

However, either negative facts are facts of a special sort, or they are just ordinary facts. In the former case, the fact that the fruit is not ripe is marked by a special glow, a quality that makes it altogether different from the fact that the fruit is ripe. It has an extra property of Negativity, or Markedness, as it is sometimes called in linguistics (see Haspelmath, 2006, for a comprehensive overview): the information conveyed by *not-A* is inherently abnormal, irregular, less natural or somehow more complex than information conveyed by *A*. Let's suppose that *not-A* is so marked.

The markedness of negation must have its source outside of language if it is to be of explanatory value: it could be mental or wordly. If the markedness of *not-A* is cognitive, we get the direction of causality backwards: the sentence *not-A* may well be more complex and harder to process than *A*, but this is because the presence of negation leads to complexity, not the other way round. After all, information that *Every student is asleep* is not more complex than information that *Some student is awake*, even though the former is equivalent to *No student is awake*. Psycholinguistic evidence has established a correlation between the use of negation and higher cognitive complexity (Dudschig et al., 2021), but this is not because negation signals a higher complexity in information that was already there to begin with, independently of our use of negation to convey it.

Alternatively, the inherent markedness of *not-A* is metaphysical, and comes from a special property of Negativity with which some facts are endowed. The nature of Negativity has long been controversial in philosophy—see the summary in Horn (1989, 50–56). Perhaps, however, Negativity is nothing too abstruse. After all there is a metaphysical contrast between presence and absence. Some animals and young children are capable of detecting this contrast and to perform basic inferences on its basis (Call, 2004; Ferrigno et al., 2021). Plausibly, our evolutionary ancestors had similar abilities. Even so, the contrast between the presence and absence of individual objects cannot be all there is to negation, since negation is not just an indication of existence vs non-existence, and perception of presence and absence of individual objects does not imply perception of presence and absence of properties. To say that Merkel is not French is not to say that Frenchness is absent when it comes to the former German chancellor. The contrast between presence and absence of objects, while familiar, could at most be a metaphysical foundation for negation in only some of its uses.

For the required generality, Negativity must encompass all negative information. But in this case, Negativity seems indeed abstruse and unfamiliar. If so, it is hard to fathom how our hunter-gatherer ancestors should have ever been concerned with such a property, being no scholars of Heidegger. That is, communication about Negativity as such could not have plausibly conferred an evolutionary advantage. Therefore, the first horn of the dilemma, on which negative information is a special property, either gets causality backwards, or undermines the case for positing benefits to communicating about it.

Perhaps the fact that the fruit is not ripe is a fact of the same kind as the fact that the fruit is ripe: negative information is just ordinary information, except that we describe it in English by the word *not*. In this case, no spooky properties like Negativity are claimed to populate the informational environment. However, on these assumptions we cannot explain why negation has the meaning that it has: why should it convey that *A* and *not-A* are incompatible.

Suppose you are looking for the keys, and I told you that the keys are in the kitchen (and I am sincere, reliable, etc.). My act of communication does not necessarily prevent you from looking for the keys in the garden. It is only because of world knowledge, according to which kitchens and

gardens are apart, that you do not go looking for the keys in the garden upon being told that they are in the kitchen. But the inference from *The keys are in the kitchen* to *The keys are not in the garden* is not forced upon us by facts alone. Compare: the fact that the keys are in the kitchen does not rule out the fact that they are on the counter. World knowledge allows for the possibility that something is both in the kitchen and on the counter. In this case, from the fact that the keys are in the kitchen we do not infer that they are not on the counter.

Furthermore, the inference from *x is P* to *x is not-Q* partly depends on matters of scope. Suppose that, as you are looking for the keys, I tell you that sometimes the keys are in the kitchen. You should not then infer that the keys are never (= not sometimes) in the garden, because *x is sometimes P* and *x is sometimes Q* are compatible subcontraries. Mere facts need not stand in relations of incompatibility with one another, and even if they do, we may easily miss this relation, for example by missing the relevant inference. Huw Price comments:

> the advantage of [negation] is that it gives us a perfectly general means of registering and pointing out the incompatibility. ...it would be useful to have a device whose function was precisely to indicate that an incompatible claim was being made ...It seems that this is what negation gives us. (Price, 1990, 224)

Thus negation expresses incompatibility, but putting bodies of information side by side does not in general result in incompatibility. Simply citing factual information as what propelled the evolution of negation, therefore, we fail to explain why *not-A* should be incompatible with *A*.

On the one hand, negative facts cannot be special facts, on pain of getting the direction of causality wrong, or of positing metaphysical powers of which it is implausible to say that we should have been communicating about over evolutionary time. On the other hand, if negative facts are just ordinary facts then there is no reason to suppose that negation would have evolved to express incompatibility. Facts-first adaptationism does not stand.

5   Frequencies-first adaptationism

There's another version of direct adaptationism which we may call 'frequencies-first adaptationism'. The view that lexicalization is often driven by frequency of use has in fact a long history in linguistics, and there is good evidence for it. Languages tend to have shorter forms for more frequently used expressions (Zipf, 1935; Piantadosi, 2014): the more an expression is used, the more speakers and listeners expect it to be used, the shorter it may become without compromising efficient communication. For example, *because (of)* derives from the phrase *by cause (of)*, which crystallized over time as the lexical item of choice to express causality in English. According to Haspelmath (2021, 614), the greater length of *not-A* is explained by its greater rarity of occurrence compared to *A*.

However, even if *not-A* is less frequent than *A*, this cannot be where explanations end. For it is also the case that past tense (*We talked*) is typically longer than present (*We talk*) and passive (*I was eaten*) is typically longer than active (*I ate*, cf. Haspelmath, 2021). However negation, past, passive, all have different forms, structures, and functions. Frequency alone, though it may be necessary, is therefore insufficient to explain negation.

To address this issue we need an argument in favor of uneven frequencies about negation—while presumably other arguments are due for past tense, the passive voice, and so on. Although this is not the point of Enguehard and Spector's (2021) paper, they do provide an argument for uneven

9

frequencies, citing psychological evidence (Chater and Oaksford, 1999). The argument is couched in Bayesian terms. Informativity is a matter of expected surprisal: the distance between likelihood of truth based on one's priors, and degree of belief upon hearing an utterance. Consider the utterance of a tautology, *Either A or not-A*. Assuming that the listener is rational, their prior degree of belief in the truth of the tautology will be 1. Upon hearing the tautology uttered, beliefs are updated by Bayes rule, resulting in a posterior degree of belief of 1. So there is no distance between prior and posterior, and no surprisal.

Suppose that the listener assumes that A is less likely to be true than *not-A*.[12] Upon hearing *not-A*, a listener will not have to revise her degrees of belief very much, with lower surprisal. Upon hearing A, Bayesian update will lead to a relatively higher surprisal. Thus A has higher surprisal value than *not-A*. Hence A is more informative than *not-A*, and so A will occur more frequently.

The argument is valid, but it crucially depends on the assumed priors. Is it in fact the case that, for rational agents, the prior probability of A is lower that of *not-A*?

> [Chater and Oaksford (1999)] observe, first, that the properties denoted by nouns, verbs and adjectives typically hold of a minority of objects (they call this observation the 'rarity assumption'): there are less cats than non-cats, and less red things than non-red things and presumably most often there are less people who are singing than people who aren't. ...There are of course obvious counterexamples (*thing*, *exist*, ...), but overall, for most lexical predicates B, fewer things have the property B than the property non-B. (Enguehard and Spector, 2021, 9)

Furthermore, Enguehard and Spector add in a footnote:

> There are several reasons why [the rarity assumption] could be true. One is that 'natural' concepts typically cover a connected and relatively homogeneous region of the space of possible concepts (Gärdenfors, 2004). To give an example, the dog-concept is arguably a more natural concept than the non-dog concept, because the concept of 'non-dog' includes many different types of objects which are intuitively extremely different from each other. (Enguehard and Spector, 2021, 9–10)

So the premise of the argument, that a rational agent's prior for A is lower than the prior for *not-A*, depends on the rarity assumption: the assumption that 'properties ...typically hold of a minority of objects' (p. 9). As Enguehard and Spector acknowledge, it is hard to assess the truth of this assumption. Moreover, there are reasons to be skeptical of it. In the next section I'll present an account of the evolution of negation that does not rely on the rarity assumption.

The rarity assumption is often false if predication is sortal, and it is inapplicable if predication is not sortal. Predication is often thought to be sortal. Sentences such as *The color of copper is forgetful* and *Friday has pneumonia* appear to be sortal violations (Thomason, 1972). In matters of speech production and interpretation, predicates come with sortal restrictions, and are understood to be relative to specific categories. If so, the rarity assumption is often false: whether A is more or less likely than *not-A* cannot be decided *a priori*, and in particular independent of what A is.

Consider the predicate *is red*, and suppose that predication is sortal. Then *is red* is well-defined only for some kinds of objects, such as flowers or berries, and not for others, such as numbers and

---

[12]What Enguehard and Spector (2021, 7) in fact assume is that the prior for *All A are B* is lower than the prior for *No A are B*, i.e., *All A are not-B*. They might reject the extension to sentences without the quantifiers. Note however that the kind of justification provided in the next citation (the rarity assumption) readily extends to the case of sentences without the quantifiers.

virtues. Then the rarity assumption is true just in case *is red* only holds of a minority of flowers, or of a minority of berries. This, in turns, depends on whether *is red* expresses a typical property of the sort, i.e., a property that probably applies to a member of the sort. This kind of typicality is often visible in some true generics, like *Roses are red* and *Cats have fur* (although the converse does not hold: many true generics express properties that do not hold for a majority of objects of sort, such as *Mosquitoes carry malaria*). Predicates expressing typical properties of members of a sort are then bound to have higher priors than their negations, hence it would be negations of such predicates that have higher surprisal value, not the other way round: *This cat has no fur* is more surprising than *This cat has fur*. The rarity assumption is that fewer things have a property than lack it. But if predicates are only defined relative to a sort, lots of predicates express properties that probably apply to all members of the sort, i.e., all the predicates that express properties that are typical of the sort. It is doubtful whether more predicates express false typical properties than true.

Perhaps predication is not sortal: in this case, a predicate is well-defined over any things whatsoever.[13] In the universal set that contains all things whatsoever, fewer things are red than not, and fewer things have fur than not. In the set of all things whatsoever the rarity assumption may well be true (except, as noted, for predicates such as *is a thing*). The problem, however, is that in ordinary language use we seldom talk about all things whatsoever, philosophers aside. But then the rarity assumption is not plausibly invoked in an evolutionary argument. We can plausibly explain how a linguistic convention evolved with respect to contexts of use that have been prominent and relevant through time. It is plausible to say, for example, that a word denoting water evolved in many languages of the world because such a word is useful in a large number of frequently occurring contexts of use. The rarity assumption seems to hold with respect to a context of use that could not have had a significant impact on our evolutionary history.

Frequency-first is insufficient to explain negation, as opposed to other constructions, without a specific argument for uneven frequencies. One such argument relies on the rarity assumption, which is true in contexts that could not have been in the driver's seat of language evolution, or else it is often false. Whether it can be defended remains a bit speculative, and an account that does not rely on this assumption is preferable, such as the account presented in the next section. Overall, frequency-first adaptationism is unconvincing.

6   Negation is not a direct adaptation

I discussed two accounts on which negation is a direct adaptation, and found both unsatisfying. Perhaps more accounts of this kind can be designed, and different arguments can be given, but I will instead outline a proposal on which negation is an indirect adaptation. Accordingly, negation is typically used to communicate that something is not the case, or that something is absent or non-existent, but an explanation of its origin does not lie in these uses. There could be many accounts of this kind, but for concreteness I'll refer to the model of Incurvati and Sbardolini (2021), on which, I argue, negation is an indirect adaptation. The authors explicitly state that negation is an adaptation, but do not explicitly address the direct/indirect distinction. Further elements of my proposal are drawn from Steinert-Threlkeld (2016), who does not address the direct/indirect distinction either.

---

[13]Enguehard and Spector (2021), p. 10, seem best understood as thinking of the rarity assumption in terms of sortal restrictions. If so, this second leg of the argument is irrelevant to their discussion. It is potentially relevant to a version of frequency-first adaptationism which rejects the idea that predicates come with sortal restrictions.

The account begins by revising some assumptions about the environment in which linguistic conventions evolved. More structure can be added to the Lewisian signalling game. It is well documented that animals can deceive, and occasionally engage in anti-social behavior for private profit: Capuchin monkeys on watch for predators sometimes send anti-predator calls only to collect all the food for themselves (Wheeler, 2009). Of course, as the boy who cried *Wolf*, excessive deception may lead to an erosion of trust, and a sentry who lies too often is quickly paid no attention. But smart monkeys can exploit the benefits of information sharing. Communication is still a cooperative game, by and large, but local infractions to cooperativity are tolerable. In this richer signalling game, the speaker communicates facts about water, fruit or predators, or about anything else in the environment, and can be truthful or not. But the listener is not completely at their mercy. The listener decides whether to accept or reject the message: to respond to the call or ignore it. If the latter, information fails to be shared and communication breaks down. But if the speaker is truthful and the listener accepts the message, the signalling game proceeds in the familiar Lewisian fashion.

Incurvati and Sbardolini (2021) add some assumptions to the basic Lewisian signalling game to account for potential conflict. First, the speaker is often truthful but need not be. The preferences of speaker and listener are not necessarily aligned. Second, the listener may but need not accept the information presented to them. The listener has a binary choice between acceptance and rejection. Both assumptions are plausibly true of many ordinary interactions, whether among humans or monkeys, and would have been so throughout our evolutionary history. This version of the signalling game with modified structure is called Rejection Game, and is represented in Figure 1. In this game, rejection is understood weakly, as a refusal to accept, rather than strongly, as acceptance of the contradictory (Incurvati and Schlöder, 2017).
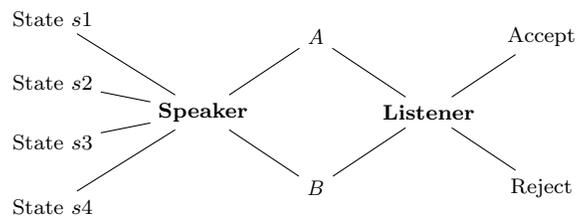

Figure 1: Decision tree of a Rejection Game with four states and two signals

In the Rejection Game, the speaker has direct evidence about which information state the interlocutors are in, among a number of possible information states (four in Figure 1), and may send a number of signals (two in Figure 1). For generality, we assume that the possible information states outnumber the possible signals on a 2:1 ratio. This, again, is quite plausible: lexical resources are finite, but the information space is potentially unbounded. (The same ratio is assumed in Steinert-Threlkeld's (2016) Negation Game, discussed below.)

Speaker and listener face a trade-off between informativity and cognitive complexity. The trade-off enforces a regime of use of signals in which synonymy (the use of more than one signal for one information state) and ambiguity (the use of one signal for more than one information state) are best avoided. The speaker cannot have a signal for each bit of information worth sharing: that would impose too heavy a toll on memory and learning. But with too few words, massive ambiguity may lead to perpetual rejection. After all, if the speaker's utterance doesn't help the listener learn their true state, the listener may decide that the speaker can be ignored anyway.

Moreover, since the states outnumber the signals, some ambiguity is inevitable in this first version of the game. More signals can be developed to distinguish between information states, but not indefinitely: the expressive resources of the language cannot grow forever. Against this background, negation is a general solution to the communicative impasse: at the small cost of a single new expression, negation allows the speaker to express in language the choice the listener has to make between acceptance and rejection. Negation evolved as a flag that can be attached to a signal to indicate that *not-A* is to be accepted if its counterpart *A* is to be rejected, and *vice versa*. This way, the speaker has a general way to avoid the ambiguity that is structurally inevitable due to communicating in a finite language about an infinite information space.

The expression of negation takes the form of a preliminary choice for the speaker, between assertion $+$ and denial $-$, who then goes on as before to choose between A and B. This preliminary choice immediately doubles the expressive power of the language, resulting in four signals in Figure 2. The speaker can now assert A or B (actions $+A$ and $+B$), or deny them ($-A$ and $-B$). The distinction between $+$ and $-$ must be marked in language to be perceived by the listener, perhaps in the form of a morpheme *not*, which indicates denial in English. In many languages, denial is so marked by negation. There are also languages, like Vietnamese (Duffield, 2007) and Coptic (Haspelmath, 2021), in which assertion is overtly marked too.
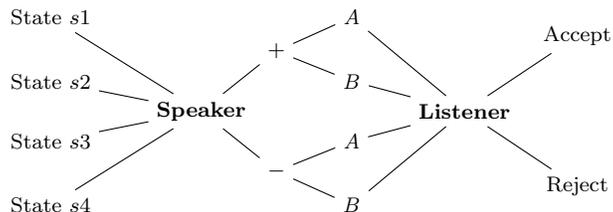

Figure 2: Decision tree of a Rejection Game with four states and two signals

Thanks to the choice between assertion and denial the speaker has the expressive power to convey information avoiding both synonymy and ambiguity. This, however, is not enough to explain negation: while the presence of a device to mark the choice is obviously preferable for communication purposes, the speaker could easily have doubled the number of available signals with any other 1-place operator. However the choice between assertion and denial is strictly tied to the listener's previously given choice between acceptance and rejection.

Acceptance and rejection are necessarily incompatible actions: one excludes the other. It is here that the incompatibility of negation goes back to. The options of accept and reject are clearly incompatible for the listener, who must pick one or the other: recall that rejection is failure of acceptance, not acceptance of the opposite. To enforce their incompatibility, repercussions for the norm-violating speaker are assumed, in the form of a specific pattern of attitudes. Thus, there are strong disincentives for the speaker using A ($+A$) and *not-A* ($-A$) interchangeably in the same state: any time the speaker so behaves, they trigger repercussions from their social environment such as scorn, mistrust or blame. This can be justified by assuming that, if a morpheme *not* has been added to A, there is an expectation of a difference between A and *not-A*, so that if these signals are then used synonymously, the expectation is violated and sanctions are imposed on those responsible. Thus the speaker has strong motivation to recognize that in a state in which A is accepted, *not-A* will be vigorously rejected, and *vice versa*, and to behave accordingly. The incompatibility in the meaning of A and *not-A* is traced to the incompatibility of acceptance and rejection.

Suppose that, playing the game over and over again, the use of A comes to be robustly correlated with state $s1$, and the use of B with $s4$.

    $s1$: Water is on the hill

    $s2$: Water is behind the mountain

    $s3$: Predators are in the forest

    $s4$: Predators are on the hill

Then there remain two states, $s2$ and $s3$, which can be expressed by *not-A* and *not-B* respectively. Negation does not mean anything, in this rudimentary language, except that A and *not-A* (and B and *not-B*) are incompatible alternatives. If one is accepted, the other cannot be. But now the states can be unambiguously sorted and information can be communicated efficiently. Hence negation comes with two novelties: it increases the expresses power of the language (so that the number of signals matches the number of states in the simple model of Figure 2) and it marks incompatibility between the use of pairs of signals A and *not-A*. This is the reason it evolves.

There remain two closely related issues. The first is that negation still has to take over the function it most often has: allowing us to convey negative information. The hypothesis is that this function is taken on only later, just as the wing first appeared and only later animals learned to use it for flying. The second issue is that, relative to the information space $s1/s4$ above, it could be for all I've said so far that A comes to mean that water is on the hill and B that predators are on the hill, while *not-A* eventually comes to be correlated with any two of the remaining states. However, it shouldn't be just random: it shouldn't be that A comes to mean that water is on the hill and *not-A* comes to mean that predators are in the forest.

We can address both issues by explaining how negation develops as a compositional operator. A model of compositional development is the Negation Game of Steinert-Threlkeld (2016).[14] Let's assume, in addition to the set of information states and the signals assumed above, that speaker and listener can recognize a relation between information states. For example, the agents can recognize a similarity relation on $s1/s4$ such that states that share the same topic (water or predators) are similar. Similarity relations are salient in our environment, and could have supported the development of a linguistic expression throughout our history. Moreover, similarity is a familiar property, not nearly as obscure as Negativity, and the assumption that our evolutionary ancestors were aware of similarity relations in their informational environment is modest. (Finally we assume, for simplicity, that the similarity relation is binary, so that each state has a unique counterpart.)

Following Steinert-Threlkeld (2016), we may account for negation as a signal − that evolves on basic reinforcement learning, and is used in combination with another signal A, to indicate the function that maps the information state designated by $+A$ to its similar counterpart. Thus, assuming that $+A$ indicates that water is on the hill ($s1$), $-A$ develops as a complex signal to indicate that water is behind the mountain, i.e., in this simplified informational context, that water

<hr />

[14]Incurvati and Sbardolini do not present a compositional model of the operator *not*, but a non-compositional model of the speech act of denial. Steinert-Threlkeld (2016) assumes that there is a relation among states that has some of the logical properties of negation, and that the speakers are aware of it. In this sense, the overall account can be interpreted as one on which negation is a direct adaptation (with the function mapping each state to its "negative" counterpart interpreted as Negativity). This, as I argued above, is not satisfactory. However, the compositional analysis of negation in Steinert-Threlkeld is a separable component of the overall account and it can be combined with Incurvati and Sbardolini's (2021) account of the source of incompatibility. My limited aim in this paper is to defend a view based on compelling aspects of both accounts.

is not on the hill. Importantly, on this account the assumed relation between information states is not contrariety or absence, but similarity (sameness of topic).[15] The contrast between A and *not-A* comes not from hypothetical negative properties found in the outdoors, so to speak, but from the opposition of acceptance and rejection. It is not from the information states themselves that negation derives its meaning of incompatibility, but from the incompatibility of acceptance and rejection.

Across languages, as noted above, there seems to be a correlation between overt morphology and frequency of use: *not-A* appears to occur less frequently than A (Haspelmath, 2021). The indirect adaptation account I sketched makes an explanation readily available. For the denial *not-A* has to be made explicit (in order to matter in the game), and that's done by adding extra morphology on the more basic signal A: by concatenating an extra symbol to it. It is enough to suppose, which seems reasonable, that some information states occur more frequently than others. Since the use of negation carries the additional cost (albeit small) of concatenation, the shorter form plausibly evolved to be used in correlation with the more frequently occurring information states, leaving *not-A* for rarer occasions of use. The relatively low frequency of negative form is therefore predicted on the basis of its communicative function, given the added cost of its employment. Uneven frequencies for the use of A and *not-A* are downstream, however, from the distinction between them.

I will conclude by summarizing the view I have presented. Negation is ordinarily used to communicate negative information. However, this function is not what drove its development. Following Incurvati and Sbardolini (2021), negation is advantageous in (mostly) cooperative communication over a large information space by means of a finite lexicon, in which a message can be accepted or rejected. The optimal solution to this interaction is one whereby the speaker has an expression to mark the difference between accept and reject, for such an expression is a general device to send a message that would be accepted in case another would be rejected. This way, the speaker has a general solution to the problem of ambiguity. (The solution is general in the sense that it's more efficient than adding one signal for each one information state, but not principled: the information space is always in principle larger than the language, and consequently not all ambiguity is ever eliminated. Languages in fact do contain ambiguity.)

Since on this game A is accepted just in case *not-A* is rejected and *vice versa*, the incompatibility between A and *not-A* is predicted. Still, the account does not yet explain how negation compositionally attaches to a signal to convey the contradictory of its argument. However, we can explain this development by assuming that speakers can recognize a similarity relation among states and then showing, following Steinert-Threlkeld (2016), that basic reinforcement learning leads to the use of *not-A* to indicate the counterpart of what's indicated by A, with the logical relation between the two signals established by the incompatibility of acceptance and rejection. Finally, we may suppose that, as the use of negation requires additional cognitive effort, more frequently occurring states are signalled by less costly signals, hence negative forms are left for relatively less frequent information.

On this account, negation is an indirect adaptation. The advantage it has and that drove its development is to allow the speaker to express the difference between acceptance and rejection. Thus the speaker can anticipate the choice of the listener, and find a general way to avoid ambiguity: if one is in a state where A will be rejected, *not-A* will be accepted, and *vice versa*. The account does not assume that negation is used to talk about negative facts, nor to convey negative information, as a cause of its development. Indeed, no assumptions are made about the information states except

---

[15]See also Franke (2016) on signalling games with a similarity metric among states used to model the emergence of compositionality.

that they stand in similarity relations with each other. Negation evolved for the speaker to stay ahead of the game of communication in a context in which the listener can accept or reject, partially overcoming the expressive limits of a finite language in a larger information space.

7 Conclusion

I have considered two versions of a direct adaptation account of negation, on which negation develops in response to a specific informational need to communicate about negative facts: facts-first and frequencies-first adaptationism. I rejected both accounts. I then presented an alternative due to Incurvati and Sbardolini (2021), and argued that on this alternative account negation is an indirect adaptation: originally developed as a general device to avoid the pitfalls of ambiguity in a game with twice as many bits of information as signals, in which the signals can be accepted or rejected. Negation is device tracks the distinction between accept and reject in language, allowing the speaker to assert and deny, and can later take on the function to express that something is not the case, or that something is absent or non-existent. The account avoids the difficulties of direct adaptation accounts of negation. Finally, I introduced aspects of the account of negation of Steinert-Threlkeld (2016), in order to explain the contradiction between A and not-A: on the final proposal, negation is a compositional signal that can evolve on basic reinforcement learning, whose meaning derives from the opposition of acceptance and rejection, and that is used to indicate incompatibility.

The are many open questions about the analogy between biological and linguistic evolution. Linguistic and biological change perhaps don't have much in common. However, just like in biological evolution, it is important to recognize a distinction in linguistic evolution between direct and indirect adaptation. Different accounts of negation illustrate this point nicely.

References

Josh Armstrong. Coordination, Triangulation, and Language Use. *Inquiry*, 59:80–112, 2016.

Mark Aronoff. Darwinism tested by the science of language. In Claire Bowern, Laurence Horn, and Raffaella Zanuttini, editors, *On looking into words (and beyond)*, pages 443–456. Berlin: Language Science Press, 2017.

Moysh Bar-Lev and Roni Katzir. Communicative stability and the typology of logical operators. ms., 2022.

Margit Bowler. Conjunction and disjunction in a language without 'and'. *Proceedings of SALT*, 24: 137, 2015.

Josep Call. Inferences about the location of food in the great apes (pan paniscus, pan troglodytes, gorilla gorilla, and pongo pygmaeus). *Journal of Comparative Psychology*, 118:232–241, 2004.

Herman Cappelen. *Fixing Language: An Essay on Conceptual Engineering*. New York, NY: Oxford University Press, 2018.

Fausto Carcassi and Giorgio Sbardolini. Assertion, denial, and the evolution of boolean operators. *Mind and Language*, 2022. forthcoming.

Nick Chater and Mike Oaksford. The probability heuristics model of syllogistic reasoning. 38(2): 191—-258, 1999.

William Croft. *Explaining language change: an evolutionary approach.* London: Longman, 2000.

Charles Darwin. *The descent of man, and selection in relation to sex.* London: John Murray, 1871.

Donald Davidson. Communication and Convention. *Synthese*, 59:3–17, 1984.

Donald Davidson. A Nice Derangement of Epitaphs. In E. Lepore, editor, *Truth and Interpretation*, pages 433–446. Oxford: Blackwell, 1986.

Donald Davidson. James Joyce and Humpty Dumpty. *Midwest Studies in Philosophy*, 16:1–12, 1991.

Kathryn Davidson. 'and' or 'or': General use coordination in asl. *Semantics and Pragmatics*, 6: 1–44, 2013.

Milica Denić, Shane Steinert-Threlkeld, and Jakub Szymanik. Indefinite pronouns optimize the simplicity/informativeness trade-off. *Cognitive Science*, 46(5):e13142, 2022.

Daniel Dennett. *Darwin's Dangerous Idea: Evolution and the Meanings of Life.* New York, NY: Simon & Schuster, 1995.

Matthew S. Dryer. Negative morphemes. In Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, editors, *The world atlas of language structures*, pages 454–457. Oxford: Oxford University Press, 2005.

Carolin Dudschig, Barbara Kaup, Mingya Liu, and Juliane Schwab. The processing of negation and polarity: An overview. *Journal of Psycholinguistic Research*, 50:1199–1213, 2021.

Nigel Duffield. Aspects of vietnamese clausal structure: Separating tense from assertion. *Linguistics*, 45:765–814, 2007. doi: 10.1515/LING.2007.023.

Émile Enguehard and Benjamin Spector. Explaining gaps in the logical lexicon of natural languages: A decision-theoretic perspective on the square of aristotle. *Semantics and Pragmatics*, 14:5, 2021.

Gareth Evans. The Causal Theory of Names. *Aristotelian Society Supplementary Volume*, 47: 187–208, 1973.

Stephen Ferrigno, Yiyun Huang, and Jessica Cantlon. Reasoning through the disjunctive syllogism in monkeys. *Psychological Science*, 32:1–9, 2021.

Jerry A. Fodor. *The Language of Thought.* Harvard University Press, 1975.

Michael Franke. The evolution of compositionality in signaling games. *Journal of Logic, Language, and Information*, 25:355–377, 2016.

Peter Gärdenfors. *Conceptual Spaces: The Geometry of Thought.* Cambridge, MA: MIT Press, 2004.

David Gil. Aristotle goes to arizona and finds a language without 'and'. In Dietmar Zaefferer, editor, *Semantic Universals and Universal Semantics*, pages 96–130. Berlin: Foris Publications, 1991.

Peter Godfrey-Smith. Three kinds of adaptationism. In S. H. Orzack and E. Sober, editors, *Adaptationism and Optimality*, pages 335–357. New York: Cambridge University Press, 2001.

Stephen J. Gould and Richard Lewontin. The spandrels of san marco and the panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society B*, 205:581–598, 1979.

Stephen J. Gould and Elizabeth S. Vbra. Exaptation – a missing term in the science of form. *Paleobiology*, 8:4–15, 1982.

Stephen Jay Gould. Not necessarily a wing. In S. J. Gould, editor, *Bully for Brontosaurus: Reflections in Natural History*, pages 199–151. London: Hutchinson Radius, 1991.

Joseph H. Greenberg. Some universals of grammar with particular reference to the order of meaningful elements. In J. Seligman, editor, *Universals of Language*, pages 58–90. Cambridge, MA: MIT Press, 1963.

Thomas Griffiths and Michael Kalish. Language evolution by iterated learning with bayesian agents. *Cognitive Science*, 31:441–480, 2007.

Patrick Hanks. *Lexical Analysis: Norms and Exploitations*. Cambridge, MA: MIT Press, 2013.

Martin Haspelmath. Against markedness (and what to replace it with). 42(1), 2006.

Martin Haspelmath. Explaining grammatical coding asymmetries: Form–frequency correspondences and predictability. *Journal of Linguistics*, 57(3):605—-633, 2021.

Laurence Horn. *A Natural History of Negation*. Chicago, IL: University of Chicago Press, 1989.

Lawrence Horn. *On the semantic properties of the logical operators in English*. PhD thesis, UCLA, 1972.

Luca Incurvati and Giorgio Sbardolini. The Evolution of Denial. *The British Journal for the Philosophy of Science*, 1:–, 2021.

Luca Incurvati and Julian J. Schlöder. Weak Rejection. *Australasian Journal of Philosophy*, 95: 741–760, 2017.

Charles Kemp and Terry Regier. Kinship categories across languages reflect general communicative principles. *Science*, 336(6084):1049–1054, 2012.

Charles Kemp, Yang Xu, and Terry Regier. Semantic Typology and Efficient Communication. *Annual Review of Linguistics*, 4:109–128, 2018.

Simon Kirby. *Function, Selection, and Innateness: the emergence of language universals*. Oxford: Oxford University Press, 1999.

Simon Kirby, Thomas Griffiths, and Kenny Smith. Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, 28:108–114, 2014.

Michael Lachmann, Szabolcs Számadó, and Carl T. Bergstrom. Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences*, 98:13189–13194, 2001.

David Lewis. *Convention*. Cambridge MA: Harvard University Press, 1969.

André Martinet. *Elements of General Linguistics*. London: Faber and Faber, 1964.

Caterina Mauri. The irreality of alternatives: Towards a typology of disjunction. *Studies in Language*, 32:22–55, 2008.

Alex Mesoudi. *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*. University of Chicago Press, 2011.

Matti Miestamo. *Standard negation: The negation of declarative verbal main clauses in a typological perspective*. Berlin: Mouton de Gruyter, 2005.

Martin A. Nowak and David C. Krakauer. The evolution of language. *Proceedings of the National Academy of Sciences*, 96:8028–8033, 1999.

Steven Piantadosi. Zipf's word frequency law in natural language: A critical review and future directions. *Psychonomic Bulletin and Review*, 21:1112–1130, 2014.

Huw Price. Why 'not'? *Mind*, 99:221–238, 1990.

Terry Regier, Paul Kay, and Naveen Khetarpal. Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, 104:1436–1441, 2007.

Mark Richard. *Meanings as Species*. Oxford University Press, 2019.

Robert Richards. The linguistic creation of man: Charles darwin, august schleicher, ernst haeckel, and the missing link in nineteenth-century evolutionary theory. In Matthias Dörries, editor, *Experimenting in Tongues*, pages 21–48. Stanford, CA: Stanford University Press, 2002.

Uli Sauerland, Ayaka Tamura, Masatoshi Koizumi, and John M. Tomlinson. Tracking down disjunction. In Daisuke Bekki, editor, *Preproceedings of Logic and Engineering of Natural Language Semantics*, volume 12, page 10. Tokyo: Ochanomizu University, 2015.

Daniel Siegel. The Failure of Condescension. *Victorian Literature and Culture*, 33:395–414, 2005.

Brian Skyrms. *Signals: Evolution, Learning, and Information*. New York: Oxford University Press, 2010.

Shane Steinert-Threlkeld. Compositional signaling in a complex world. *Journal of Logic, Language and Information*, 25:379–397, 2016.

Shane Steinert-Threlkeld. Quantifiers in natural language optimize the simplicity/informativeness trade-off. In *Proceedings of the 22nd Amsterdam colloquium*, volume 30, pages 513–522, 2020.

Rich Thomason. A semantic theory of sortal incorrectness. *Journal of Philosophical Logic*, 1:209–258, 1972.

M. Tomasello. *Constructing a Language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press, 2003.

Elizabeth Closs Traugott. Exaptation and grammaticalization. In Minoji Akimoto, editor, *Linguistic Studies Based on Corpora*, pages 133–156. Tokyo: Hituzi Syobo Publishing Co., 2004.

Wataru Uegaki. The informativeness/complexity trade-off in the domain of Boolean connectives. *Linguistic Inquiry*, 1:1–39, 2022.

Brandon C. Wheeler. Monkeys crying wolf? Tufted capuchin monkeys use anti-predator calls to usurp resources from conspecifics. *Proceedings of the Royal Society B: Biological Sciences*, 276: 3013–3018, 2009.

Amotz Zahavi. Mate selection—a selection for a handicap. *Journal of Theoretical Biology*, 53: 205–214, 1975.

Noga Zaslavsky, Mora Maldonado, and Jennifer Culbertson. Let's talk (efficiently) about us: Person systems achieve near-optimal compression. In *Proceedings of the 43rd annual meeting of the cognitive science society*, 2021.

George Kingsley Zipf. *The Psycho-biology of Language: An Introduction to Dynamic Philology.* Cambridge, MA: MIT Press, 1935.

George Kingsley Zipf. *Human Behavior and the Principle of Least Effort.* Oxford: Addison-Wesley Press, 1949.