




The Evolution of Inclusive Folk-Biological Labels and the Cultural Maintenance of Meaning

Ze Hong^{1,2} 

Accepted: 22 March 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

How is word meaning established, and how do individuals acquire it? What ensures the uniform understanding of word meaning in a linguistic community? In this paper I draw from cultural attraction theory and use folk biology as an example domain and address these questions by treating meaning acquisition as an inferential process. I show that significant variation exists in how individuals understand the meaning of inclusive biological labels such as “plant” and “animal” due to variation in their salience in contemporary ethnic minority groups in southwest China, and I present historical textual evidence that the meaning of inclusive terms is often unstable but can be sustained by such cultural institutions as religion and education, which provide situations in which the meaning of linguistic labels can be unambiguously inferred.

Keywords Language evolution · Semantic change · Cultural attraction theory · Folk biology

One extraordinary feature of human communication is that we are capable of representing concrete reality using rather abstract symbols, which greatly facilitates and shapes the cultural transmission of information (Castro et al., 2004; Gelman & Roberts, 2017). But how do human individuals acquire the meaning of these rather arbitrary symbols? Philosophers have long noticed this problem: early empiricists such as Locke (1689/1847) and Hume (1739/2003) indirectly treat it as examples of knowledge acquisition via associative learning, and later thinkers such as Wittgenstein (1953/2009) and Quine (1960/2013) explicitly lay out the inferential problem of attributing meaning to linguistic labels. Briefly, the puzzle is how one can

✉ Ze Hong
ze_hong@g.harvard.edu

¹ Department of Sociology, Zhejiang University, Yuhangtang Road 866, Hangzhou 310058, China

² Department of Human Evolutionary Biology, Harvard University, 11 Divinity Avenue, Cambridge, MA 02138, USA

be certain that a particular linguistic label denotes some concept from observing the label being invoked in specific occasions since there are an infinite number of concept-label matching possibilities. To borrow Quine's famous example: observing a rabbit scurry by and the native says "*gavagai*," a foreign linguist can never be fully certain of its meaning. The word "*gavagai*" could refer to rabbits, or any mammal, or any animal, or any object, or even white or furriness (Bloom, 2002)!

Considerable work in contemporary psychology and psycholinguistics has been devoted to this issue. Today, word-meaning acquisition is often described as an inferential task which is solved through hypothesis testing or Bayesian updating (Xu & Tenenbaum, 2007). Specifically, learning the meaning of words involves determining the mapping between linguistic labels and mental concepts via experience. However, there are heavy constraints regarding the initial hypothesis space; in other words, humans have prior knowledge about what possible candidate hypotheses are (Carey, 1978). Without such priors, word learning would be computationally impossible.¹

Much linguistic work, however, presumes that humans (at least adults) somehow successfully solve the problem of word-meaning acquisition, with the implication that individuals within a linguistic community agree on the meaning of linguistic labels. Of course, this is often acknowledged as a simplifying assumption, and the extent to which internal representations of meaning actually align in communities is an open question. It is quite possible that nonuniform understanding of word meaning exists in many domains (Geeraerts et al., 2010), and my aim in this paper is to explicitly address such non-uniformity by highlighting the role of cultural institutions in regulating meaning attribution. In particular, I use folk biology as an example domain to address two questions: (1) Is it possible for individuals in a population to attribute different meanings to the same linguistic label? If so, in what way? (2) Why do certain societies have inclusive linguistic labels such as "*plant*" and "*animal*" while others lack them? I suggest that the two seemingly unrelated questions can be answered by invoking the same psychological and social processes. To do so, I draw from research in cultural evolution, in particular cultural attraction theory (Buskell, 2017; Scott-Phillips et al., 2018; Sperber, 1996), and argue that certain inclusive linguistic labels are unstable with regard to their meaning,² and cultural institutions such as mass education and religion often serve as maintainers of word meaning which ensure both the existence of inclusive linguistic labels and the uniformity in attributing meanings to these labels.

In the rest of the paper, I will first briefly summarize the existing literature in folk biology as well as the relevance of cultural attraction theory and then point out the important role that cultural institutions play in shaping and maintaining the meaning

¹ The focus of this paper is on lexical acquisition rather than concept acquisition. Although my thesis is most compatible with the nativist view of concepts, it does not depend on radical concept nativism in the style of Fodor (1975, 1983) as long as there is substantial uniformity in understanding of biological concepts. Such uniformity in the possession of concepts could be due either to their innateness or that they are acquired through some learning mechanism in a shared environment.

² Throughout this paper, *stability* will be used to refer to uniform understanding of the meaning of a given linguistic label.

of folk-biological labels. I then present rather different types of evidence to support my arguments: individual-level meaning attribution data from 23 ethnolinguistic groups in southwest China, which includes a quasi-natural experiment regarding the presence/absence of cultural institutions among the Lisu people in Yunnan province, and a historical textual analysis on the semantic evolution of the folk-biological labels in Mandarin Chinese. These data collectively show that the meaning of the more inclusive labels (e.g., *plant* and *animal*) often vary across populations and throughout history, and their stable existence is usually the result of top-down institutional influences.

Concepts and Labels in Folk Biology

Folk biology has been one of the major domains (along with color and kinship) for cross-cultural comparisons in categorization and linguistic labeling. This is because every human population has some experience with plants and animals, and given the natural taxonomic structure of living organisms, folk biology serves as an ideal candidate for evaluating claims on human cognitive universals and cultural variations. In his seminal paper, Berlin et al. (1973) proposed some general principles of classification in folk biology: living organisms are classified into five ranks from the most inclusive to the least inclusive: unique beginner, life-form, generic, specific, and varietal. Here, “unique beginner” refers to the most inclusive categories, e.g., *plant* or *animal*.³ “Life-form” refers to the level that is one rank lower, usually with some prominent perceptual features; examples include *tree* and *bird*. “Generic” is roughly equivalent to the Western concept of species and is the most numerous and psychologically salient. The last two ranks are more specific and often contain more than one semantic dimension (e.g., *red rose*).

Berlin points out that unique beginners are often not explicitly labeled in non-Western, small-scale societies. Note that the lack of linguistic label does not mean lack of mental concept; in fact, there are so many “recognized” but unnamed categories that Berlin et al. (1968) created the “covert taxa” category for them. These categories are “covert” because natives often are able to state that species *x* and species *y* belong together in that they are “companions” and sort organisms into somewhat stable categories in pile-sorting tasks. Relating back to unique beginners, this suggests that despite the lack of explicit labels, humans everywhere may nonetheless be able to distinguish the categories *plant* and *animal* conceptually. Some recent developmental studies support such universality, showing that infants as young as 8 months old can distinguish animals from other objects and have certain expectations about them (e.g., that they have filled insides; see Setoh et al., 2013).

Evolutionarily minded anthropologists such as Atran (1998) take a strong view that the human mind has evolved hierarchical ranks that correspond to the natural

³ Curiously, Berlin et al. (1973) do not explicitly state that the scientific concepts of *plant* and *animal* are the only unique beginners. In principle, any category that is immediately above life-form should qualify (Urban, 2010); however, to my knowledge, no other unique beginners have been proposed.

grouping of the biological world. Such mental organization enables crucial cognitive functions such as category-based induction. For example, Atran posits that humans everywhere share the same ranking system, which includes the category level that represents the most fundamental division of all living organisms, *plant* and *animal*.⁴ Knowing that a particular object belongs to the category “animal” thus enables us to make a series of inferences regarding its physical properties and behaviors. As far as inductive reasoning is concerned, whether there are explicit linguistic labels is of secondary importance so long as these mental categories exist.⁵

Naming practices, however, have been investigated along a different line of research. Berlin (1977), for example, speculates that from the perspective of historical development, generics are the first to appear in any folk-biological nomenclature system and unique beginners, on the other hand, are the last. An implication of this view is that societies at certain stages of development do not “need” these inclusive categories: “while man has no doubt tacitly recognized the world of plants as a conceptual category since earliest times, it does not appear to have been essential to provide the concept with a distinctive label until quite recently” (Berlin, 1977).

Similarly, Brown et al. (1985) suggest that vocabulary of the biological world largely reflects the long-term interest of people in specific cultural settings. Agriculturalists have larger biological taxonomies than hunter-gatherers because of the cultivation and domestication of wild organisms. Hunn (1982) emphatically points out that a folk classification system has a strong utilitarian component; that is, concrete and specific categories exist because they are pragmatically useful. For example, the Tzeltal exhibit very little interest in butterflies and moths but have a detailed terminal folk taxa for their larvae, which are of practical significance: some are edible, others attack crops, etc. (Hunn, 1977). To summarize, anthropologists generally agree that while people possess the conceptual categories at each rank level, linguistic labels may be missing for culture-specific reasons.

Folk-Biological Concepts as Cultural Attractors

First proposed by the French cognitive theorist Dan Sperber (1996), cultural attraction theory (CAT) is a framework to explain the transmission and transformation of cultural information such as ideas, beliefs, and values. In contrast to another school of thought in the field of cultural evolution which takes learning or the acquisition of cultural variants to be replicative (Boyd & Richerson, 1985; Henrich & Boyd, 2002), CAT emphasizes the reconstructive nature of information transmission which relies on humans’ inferential capacities. When an individual acquires some new cultural item, he does not simply copy the variant as is; rather, he constructs a variant of his own based on background knowledge, inferential capacities, and so on

⁴ Atran seems to suggest that *plant* and *animal* are the only possible unique beginners.

⁵ Whether folk biology deserves its own ontological category is a subject of debate. Susan Carey, for example, suggests that folk-biological cognition is embedded in a more general explanatory framework of folk psychology, and that a proper understanding of the living world requires some conceptual change (Carey, 1985; Carey & Spelke, 1994). See Vapnarsky et al. (2001) for a response.

(Claidière & Sperber, 2007). During such a reconstructive process, certain outcomes are more likely to occur than others. In Sperber's own words: "Why should there be cultural attractors at all? Because there are in our minds, our bodies, and our environment biasing factors that affect the way we interpret and re-produce ideas and behaviors. . . . When these biasing factors are shared in a population, cultural attractors emerge" (Sperber, 2012). In the cultural epidemiology literature, these biasing factors are referred to as "factors of attraction" (Scott-Phillips et al., 2018).

How does this relate to linguistic labels? Since meaning acquisition can be viewed as associating labels with concepts (Clark, 2017; Macnamara, 1982), and a significant amount of language learning occurs in informal settings where no word definitions are provided (Callanan et al., 2011), the learner has to infer the meaning of the word x by observing x being used in specific occasions. If we treat word learning as hypothesis testing or Bayesian updating, the "hypotheses" or "priors" may be viewed as factors of attraction in the sense that the learner is more likely to infer particular label-concept matchings than others. Note that the data to be presented in this paper are not empirical tests of CAT per se; rather, CAT is invoked here as an explanatory framework that provides useful conceptual tools for better understanding the stability of cultural items over time and space (Heintz, 2018) and, as such, can help us think and talk about the regularities in human inference and learning.

In folk biology, different ranks have different levels of salience. Based on extensive ethnographic work, Berlin concluded that the most salient rank is folk generics (roughly corresponding to scientific genera or species; see Berlin, 1977, 1992), which has been supported by subsequent anthropological work (Atran, 1999; Brown et al., 1986). My own fieldwork in southwest China also shows that indigenous people almost always use folk generic terms to answer the question "What is this?" when pictures of animals and plants are presented (unpublished data). Research in cognitive psychology provided further evidence for the existence of a psychologically salient or "basic" rank, although the location of this rank depends on the cultural background of the subjects (Coley et al., 1997; Rosch et al., 1976). For people in small-scale, traditional societies, such as the Maya in lowland Guatemala, the basic ranks are indeed folk generics; for undergraduate students at Berkeley, however, the most privileged rank seems to be life-form (e.g., *fish*, *tree*).⁶ In either case, some rank(s) is (are) more salient than others, but unique beginner terms are never very salient comparatively. In language learning, this way of narrowing the hypothesis space by assuming that a linguistic label extends to all and only members of a particular ontological kind is sometimes referred to as the *taxonomic assumption* (Markman & Hutchinson, 1984). In the framework of CAT, the association between linguistic labels and rank-level concepts would be factors of attraction, and the unique beginner-referent matching may be a particularly weak one (Fig. 1).

Of course, an individual may encounter multiple occasions when an inclusive label is used to refer to very different organisms so that an inclusive meaning can be

⁶ Coley et al. (1997) suggest that the shift in basic rank for American undergraduates is due to a lack of specific knowledge about the biological world.

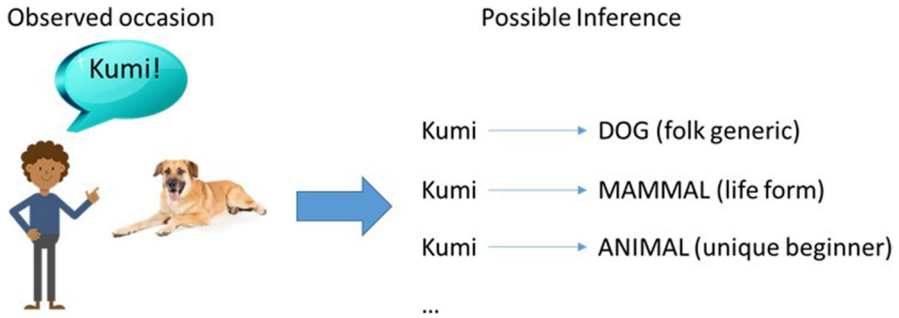


Fig. 1 Graphic illustration of folk-biological concepts as cultural attractors in an observational learning setting. The word “*kumi*” may refer to multiple concepts, and the occasion here does not definitively determine which concept the speaker has in mind

inferred. This, however, depends on the idiosyncratic experience of the individual and opens up the possibility of nonuniform understanding of the same linguistic label. If we temporarily suspend the assumption that there is some “true meaning” out there to be grabbed by naive learners, then it would be quite extraordinary that everyone in a community shares the exact same label-concept matching. Again, in the language of CAT, the coexistence of multiple factors of attraction is likely to result in a distribution of understandings based on the salience level of different biological ranks. For example, at a given time, 50% of the people in a linguistic community may think the word “*kumi*” refers to the concept dog, 30% think it refers to mammal, and 20% think it refers to animal. If the variance of the distribution is small, a naive anthropologist is likely to miss the variation or treats it merely as noise. She may conclude that the unique beginner terms do not exist in a population when in fact a small number of individuals do think that some word refers to the inclusive concept animal or plant. On the other hand, if most individuals in a population possess such matching, then the anthropologist may conclude (in a sense, justifiably) that linguistic labels that refer to unique beginner concepts exist.

To what extent is there such variation in individuals’ understanding of the meaning of linguistic labels? This question is particularly tricky in small-scale, preliterate societies where no dictionary exists. Such variation may be interpreted as certain individuals not possessing the “correct” understanding of words, yet linguists have long abandoned the idea that there is some “cultural truth” regarding the meaning of words and (sociolinguists in particular) instead have focused on the entire distribution of linguistic understanding and behavior in a linguistic community (Roberts & Sneller, 2020), as well as the culture-specific nature of lexical semantics (Malt & Majid, 2013). In the following section, I will follow this line of work and show how sociocultural institutions modulate the salience of different biological ranks and in particular how they reduce the variation in individuals’ understanding of the unique beginner labels in folk biology.

The Maintenance of Meaning by Cultural Institutions

In modern, literate societies, dictionaries and language academies sometimes claim to be the arbiter of truth on word meaning and often in effect play a role in regulating language use. Standardization of indigenous languages also occurs in contemporary small-scale societies (Andronis, 2003; Romero, 2012), but such efforts are relatively recent and often incomplete. In these preliterate linguistic communities, no one is policing the use of words and lexical acquisition occurs primarily in informal contexts. As a result, there are few occasions in which one can unambiguously map a given linguistic label to the concept animal or plant. When people in these societies talk about living organisms in everyday conversations, they very rarely invoke the inclusive concept animal or plant. In southwest China, most of local peasants' conversations involving living organisms are on folk generic species and are often about attributing some general characteristics to individual species: wild chickens have beautiful feathers, bears are dangerous, certain spiders are inedible, etc. No one seems to bother to explicitly theorize on the deep commonalities between bears and spiders in everyday situations (Hong, unpublished data).

In such an environment, the meaning of words may be considered *unregulated*. Effective regulation of word meaning, however, can happen in the presence of institutions. In the case of folk biology, these institutions can alter the salience level of different biological ranks (e.g., increase the salience of the more inclusive ranks) by providing word learning settings where word meaning can be unambiguously inferred or obtained. One obvious institution that maintains and homogenizes individuals' understanding of inclusive labels is mass education—in particular, formal biology classes in the case of learning the words that represent living organisms. For example, in a modern classroom setting a biology teacher may offer explicit verbal instruction to students: “*Cats, dogs, and cows are mammals. Sparrows and crows are birds. They are all animals.*” Such a situation makes clear that the word *animal* cannot only refer to mammal or bird, or any folk generic species but the more inclusive concept animal, which the students presumably already possess. Although individuals are psychologically prepared to prioritize folk generics (or in some cases, life-form), formal biological education provides unambiguous situations in which one can consistently and reliably attribute the linguistic label to the more inclusive concept.

Another, perhaps less obvious mechanism is religious texts that involve creation stories. All Abrahamic religions, for example, have sacred texts that vividly describe how God creates the world and living things. In the Old Testament, we see the following passage:

And God said: Let the earth bring forth the living creature in its kind, cattle and creeping things, and beasts of the earth, according to their kinds. And it was so done. (Genesis 1:24)

Here, the phrase “living creature” is used. Note that this is the King James Version, which first appeared in 1611. In the Latin translation of the Bible (Vulgate), which has been the official script of the Catholic Church since the fourth century, the word used is *animam*:

dixit quoque Deus producat terra animam viventem in genere suo iumenta et reptilia et bestias terrae secundum species suas factumque est ita. (Genesis 1:24)

In Latin, “animam” (accusative singular of “anima”: soul, spirit, breath) clearly shares the same etymological root with “animalia” (animal, living creature), whose variants gradually became the linguistic label for the concept animal in many European languages. Similarly, in the Quran of Islam, there are also passages invoking the general concept of the living creatures:

This life of the world (الْحَيَاةِ الدُّنْيَا) is but a pastime and a game. Lo! The home of the Hereafter – that is Life, if they but knew. (Surah Al-‘Ankabut 29:64)

In this passage, the Arabic word الْحَيَاةِ الدُّنْيَا (*hayawan*) is used to denote living organisms in general, which is used in many Muslim countries as a loanword. Another frequently occurring term in Quran that has a rather inclusive meaning is *dabba* (دَابَّةٌ), which usually gets translated as “moving/living creature” in English. For example:

He has created the skies without any supports that you could see, and has placed firm mountains upon the earth, lest it sway with you, and has caused all manner of living creatures (دَابَّةً) to multiply thereon. And we send down water from the skies, and thus we cause every noble kind [of life] to grow on earth. (Surah Luqman 31:10)

Here, the contexts in which these inclusive terms appear are also creation stories. Such religious texts thus serve as important cultural support for word learning by providing clear situations in which the inclusive meaning of a word can be unambiguously inferred.

Nonuniform Meaning Attribution: Evidence from 23 Ethnolinguistic Communities in Southwest China

To empirically examine the extent to which individuals may have different meaning attributions of biological labels, in the summer of 2018 I conducted fieldwork on folk-biological labeling among multiple ethnolinguistic groups in southwest China, including Mandarin-speaking individuals in Ning’er and Shuangjiang in Yunnan Province (Fig. 2). Although the Chinese government officially recognizes only 55 ethnic minority groups, there are many more subgroups with unique local dialects. For example, more than 9 million Yi people reside in Sichuan, Yunnan, Guizhou, and Guangxi provinces, and at least four of the dialects are mutually unintelligible (Bradley, 2004). For my purposes, two local dialects are deemed as different languages if (1) most (> 50%) informants in both communities agree that they are different dialects⁷ and (2) most (> 50%) informants in both communities cannot recognize the pronunciation of words for three major domestic animals (*chicken*, *pig*, *cow*) as well as *bird* and *fish*, or the most inclusive term for *animal* of the other dialect.

⁷ Informants were asked the following question “Is X people’s dialect the same as Y people’s?” and were forced to give a yes/no answer.



Fig. 2 Map of ethnolinguistic groups in Southwest China visited during this research. Locations of ethnic minority groups are noted in parentheses

In each ethnolinguistic community, I first employed convenience sampling to solicit names of locally knowledgeable individuals (i.e., asking people “who are knowledgeable regarding mastery of the local language”), whom I then interviewed to obtain a list of folk-biological labels of various levels of inclusivity. The interviews were conducted in the local language with the help of a translator (who is proficient in both Mandarin and the local language), and a hierarchical ranking was

subsequently constructed. In translating these biological labels, we relied heavily on the translators' judgement regarding how to best match the local terms with a Mandarin label (or multiple labels), which I then translated into English. Consistent with classic findings in folk biology, folk generics are the most numerous while labels for life-forms are few, and terms for unique beginners are almost always lacking (Berlin et al., 1973).

To elicit label-concept matching at the individual level, 10–25 people in each community were shown a picture of a typical local animal and asked the following question “Does this belong to X?” with X being the linguistic label for folk generic, life-form, or unique beginner (if present).⁸ Each picture was presented multiple times to obtain as many labels at different ranks as possible. To investigate the extent to which the same linguistic label may be understood differently, I analyze how individuals attribute folk generic species to the most inclusive label within a linguistic community. In some communities, domestic animals form their own category, in which case the term that roughly translates as “wild animal” is examined. Following Brown (1984), I classify wild living organisms into the following conceptual categories based on natural discontinuities at the life-form level: mammal, bird, and wug (worm + bug).⁹ Table 1 shows the detailed breakdown of the types of animals used. Note that both mouse and snake are grouped into wug because the native species are relatively small.

The results are shown in Fig. 3 with some remarkable patterns. First, although all subjects agree on the names of specific animals at the folk generic level, substantial variation regarding whether they belong to the most inclusive label exists in all linguistic communities, except for Mandarin. Coincidentally, the most inclusive label in Mandarin, *dongwu* (動物),¹⁰ also includes all animals presented, thus truly representing the concept animal. Second, the variation in individuals' meaning attribution is very systematic and seems to occur at the life-form level; that is, individuals tend to think either an entire life-form class belongs to the most inclusive label or it does not (except for mouse, bat, and snake, which are ambiguous at the life-form level). For example, among the Yi in Puxiongzhhen, Sichuan, 42% of the subjects believe that the most inclusive term “*nibuhibu*” includes all animals, whereas 33% of the subjects believe *nibuhibu* includes mammal and bird, but not wug, and 25% of the subject believe that *nibuhibu* only includes mammal. Most notably, none of the subjects think that *nibuhibu* includes mammal and wug, but not bird. In other words, there seems to be a hierarchy of the kind of life-forms that would be included in the inclusive label. Everyone agrees the most inclusive label includes large quadrupeds such as wild pig and bear. Some think that bugs and worms should be excluded; of these people, some further believe that flying, feathered creatures should also be excluded.

⁸ As will be seen, most linguistic communities do not have unique beginner terms, in which case we tried to elicit the terms that are more inclusive than the life-form level. These more inclusive terms may contain one or more life-forms.

⁹ The term wug here follows from Brown's (1984) proposed classification scheme and is not related to “wug test” in psycholinguistics (Berko, 1958).

¹⁰ Traditional Chinese characters will be used throughout this paper.

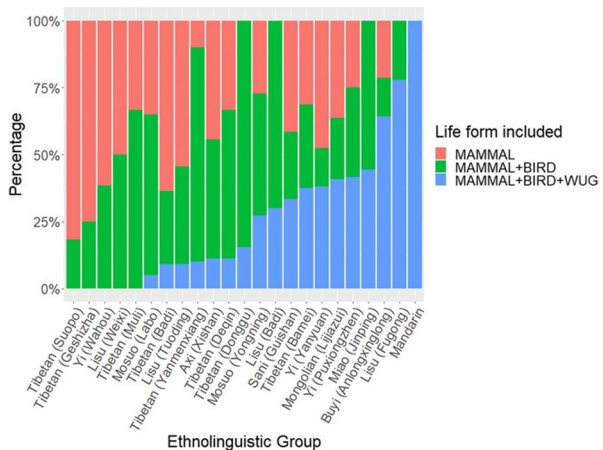
Table 1 Animals used in classification task

MAMMAL	BIRD	WUG
sheep	wild fowl	ant
panda	magpie	butterfly
wolf	kite	housefly
beaver	owl	dragonfly
boar		mosquito
bear		earthworm
		bee
		beetle
		frog
		snake
		mouse

Subjects were presented with pictures of individual animal species and were asked to indicate whether the animal belongs to some life-form or unique beginner

The mere presence of individual-level variation in how people understand the meaning of linguistic labels is not surprising; what is notable here is that the variation is nonrandom, and the pattern is highly suggestive of substantial variation of salience of different biological ranks (or, in the language of CAT, the existence of multiple factors of attraction). When individuals encounter situations in which they need to infer the meaning of some inclusive label for living organisms, there are strong inductive biases that make attributing an entire conceptual category (e.g., mammal, bird, wug) to the label much more likely than arbitrary combinations of individual folk generic species. In most communities, the salience of the most inclusive rank is not strong enough to induce convergence in meaning. So, what is special about Mandarin? Obviously, as the official language of China, Mandarin is different from the rest of the indigenous languages in many ways: it has a well-established

Fig. 3 Variation in individuals’ meaning attribution of the most inclusive label across 24 ethnolinguistic communities. Horizontal axis represents different ethnic groups, with location indicated in parenthesis. As mentioned in the main text, two groups may have the same officially assigned name yet people in these groups speak mutually unintelligible dialects. Sample size in each group ranges from 10 to 25 individuals. Mandarin data were collected in Ning’er and Shuangjiang in Yunnan Province



written form with a long literary tradition; it is the lingua franca of ethnic groups in southwest China,¹¹ and being able to speak it is often a signal of status and power. Regarding the inclusive nature of the unique beginner term *dongwu* and the uniformity of the understanding of its meaning, I suggest that formal biological education plays a crucial role. As mentioned, many people explicitly state that they learned the word *dongwu* in school, and younger informants would frequently try to explain the meaning of *dongwu* by invoking the *plant* vs. *animal* distinction, both of which were likely acquired in a formal educational setting.

What about religious institutions? As mentioned, religious institutions also provide learning contexts in which the meaning of inclusive folk-biological labels may be unambiguously inferred. Here I refer to my own fieldwork among the Lisu in southwest China as a natural experiment. Lisu people in Fugong County have a long history of practicing Christianity imported by missionaries during the 1920s, who also created written forms for the local language using the Roman alphabet (Liu, 2019) and translated the New Testament. During my visit in 2018, Christianity was prospering in Fugong County,¹² and data from the early 2000s suggest that more than 60% of the total population in Fugong are active Christians (Ying, 2009). Interestingly, Liukuzhen, about 130 km from Fugong, is much less religious, possibly due to its proximity to the capital of Nujiang Autonomous Prefecture and thus being more influenced by mainstream Han culture. As I interviewed people in Fugong, they often suggested that I “check the Bible” for animal terms, and I learned that the term “*huashihuazhi*,”¹³ which literally translates as “all kinds of meat,” was created by early missionaries and frequently used to mean the concept animal. To get a better sense of how individuals in these two localities attribute meaning to this term, I conveniently sampled 22 interviewees in Fugong and 18 interviewees in Liukuzhen and asked them, “Do you recognize the word *huashihuazhi*” and when the answer was yes, a series of follow-up questions—“Does Y belong to *huashihuazhi*?” (where Y is a folk generic term from Table 1)—to elicit their understanding of this term. In Fugong, 18 of the 22 (81.8%) people interviewed recognized “*huashihuazhi*” and 14 individuals attributed the most inclusive animal concept to it. In contrast, in Liukuzhen, where Christian churches and religious texts are largely absent, only 2 of 18 (11.1%) recognized this word. Most individuals in Liukuzhen stated that they have never heard of this word and find the expression odd.

The fact that the two ethnolinguistic groups (Mandarin and Lisu in Fugong) with the most individuals possessing the most inclusive folk-biological label are also heavily influenced by formal educational and religious institutions highlights the stabilizing role of these institutions. (The Lisu case is particularly illuminating in that the word for the unique beginner was intentionally created to represent a concept

¹¹ Strictly speaking, most ethnic groups in this area use local versions of southwestern Mandarin (西南官话), which is phonetically distinct from standard Mandarin (普通话) but is generally intelligible to standard Mandarin speakers.

¹² In the Chinese provincial system, “county” (县) is a larger geopolitical unit (in both size and population) than “*zhen*” (镇).

¹³ The Lisu Bible has other expressions for the inclusive concept animal, such as “*huazabieza*.” Another term for all sentient beings is “*sashisazhi*.”

that already existed but was not communicated often enough for a stable linguistic label to emerge.) Individuals in a community do not magically associate a given linguistic label with the same concept: uniformity in understanding lexical meaning depends on the uniformity of individuals' cognitive interaction with the external world (Gasparri & Marconi, 2019), and there is no a priori guarantee that individuals' idiosyncratic experiences will lead them to converge on the same meaning for any linguistic label.

To further investigate how the tendency for individuals to include more folk generic concepts in the most inclusive label they possess correlates with individual-level demographic variables, I performed a mixed effect regression model (with ethnolinguistic group as the random effect) that includes sex and age, as well as whether the folk-biological labeling is influenced by religion or education (i.e., 1 for Lisu in Fugong and Mandarin individuals, 0 for all other individuals), and the regression results can be seen in Table 2. As can be seen, whether an individual's biological labeling is affected by religion or education is the only variable that significantly predicts their tendency to have more inclusive labels.

It should be acknowledged that the observed individual-level heterogeneity is open to different interpretations. One obvious explanation is that inclusive labels in folk biology are often polysemous (same word with multiple related meanings), and the observed variation in meaning attribution simply reflects variation in subjects' lexical competence. This explanation depends on the existence of some "lexical truth" which, given the conventional nature of language, is difficult to establish objectively. In many ethnic groups in southwest China, knowledgeable individuals,

Table 2 Regression results for the inclusiveness of the most inclusive label that individuals possess over demographic variables

Predictors	Inclusive score			
	Estimates	CI	<i>p</i>	<i>df</i>
(Intercept)	2.03	1.62, 2.45	<0.001	137.84
sex dummy	-0.11	-0.32, 0.10	0.300	177.20
education	-0.01	-0.04, 0.01	0.333	186.99
age	-0.00	-0.01, 0.01	0.717	185.63
religion or education	1.01	0.27, 1.75	0.010	20.72
Random Effects				
σ^2	0.41			
τ_{00} language	0.14			
ICC	0.25			
N_{language}	23			
Observations	192			
Marginal R ² / Conditional R ²	0.140 / 0.357			

"Inclusive score" denotes recoded life-form-level concepts that the most inclusive label contains: mammal = 1; mammal + bird = 2; mammal + bird + wug = 3. As such, the higher the value, the individual's most inclusive label contains more life-form concepts. Regression analysis was performed using "lme4" package and visualized in a tabular format with "sjPlot" package in r. Bold values are significant at the $p < 0.05$ level

the supposed lexical experts, very frequently disagreed about whether specific folk generic species should be included in some general linguistic label and would engage in passionate debate with each other.¹⁴ My data show that older people are no more likely to possess inclusive labels ($p=0.717$). Therefore, people do not seem to gradually become “competent” as they age. Another explanation is that these ethnic minority groups are undergoing cultural and linguistic transition as younger individuals are increasingly receiving formal education and forgetting their native language and older individuals are also influenced by economic and cultural modernization. The hypothesis would be that as people become more educated in Mandarin and have more opportunity to encounter the inclusive concept animal in mainstream Han culture, they will associate this inclusive concept with the linguistic label in their native language; in other words, less “traditional” individuals will think the same label is more inclusive than traditional individuals do. My data cannot definitively rule out this alternative, and further research in small-scale societies that are relatively unaffected by external cultural influence would be needed. Preliminary regression analysis in Table 2, however, shows that the educational level as measured by years of schooling does not significantly correlate with inclusive score, indicating that schooling itself may not be a potent enough cultural force to shape individuals’ lexical understanding.

It is certainly true that cultural changes are happening, yet this should not be the reason to trivialize the within-group variation observed here. First, it is unlikely to explain all the variation observed. More importantly, the variations are not random; the observed patterns nicely illustrate the inductive biases people use in making inferences in the domain of folk biology. All subjects in Fig. 2 completely agree on linguistic label designation at the folk generic level, and most agree at the life-form level,¹⁵ and it is only at the unique beginner level that people’s understanding starts to diverge. This suggests that the unique beginner labels are the most unstable and would be the first to lose their inclusive meaning during the process of linguistic change.

The ethnographic data used here should be viewed as suggestive owing to limited sample size and nonrandom sampling,¹⁶ and further studies are needed to account for additional patterns (e.g., why closely related ethnolinguistic communities do not seem to have similar patterns).¹⁷ They do, however, provide corroborating evidence for the thesis that nonuniform understanding of some inclusive linguistic labels may be an expected feature of human language and therefore deserve more scholarly attention. In

¹⁴ In fact, polysemy itself has been suggested to be a by-product of semantic change by historical semanticists (Bréal, 1904), with the implication that polysemy may be a sign of meaning instability.

¹⁵ The main disagreements at the life-form level are whether “hawk” belongs to the inclusive category “bird.”

¹⁶ Due to logistic constraints and the fact that some of the indigenous languages are no longer spoken by younger individuals, we often asked local people to “recommend” individuals who are knowledgeable in the local language.

¹⁷ As a result of the region’s complex migration and relocation histories, ethnolinguistic groups that are geographically close together (or even officially designated the same group) often speak quite different languages. For example, the four Tibetan groups in Danba (the northernmost groups in Fig. 2) all speak their own dialects (which have little to do with the standard “Tibetan” spoken in Tibet) and are mutually unintelligible.

the next section I shift my focus to textual evidence from ancient languages, using classic Chinese as an example to further illustrate the instability of unique beginner terms.

Historical Textual Evidence on the Instability of Inclusive Biological Labels

One major theory in linguistic anthropology to explain the lack of inclusive biological labels is that these inclusive terms depend on societal complexity (Brown et al., 1985): as societies increase in scale, inclusive terms become more psychologically salient and thus are more likely to be maintained. In the language of CAT, large-scale societies can be said to somehow strengthen the factors of attraction of the linguistic matching between labels and inclusive concepts. In the case of unique beginners, Urban (2010) surveyed a large number of languages and shows that societies that rely on hunting and gathering are more likely to lack these terms than those relying on other subsistent types. The existing work, however, lacks a mechanism. It does not specify exactly how larger-scale societies achieve this, or the kind of situations larger-scale societies provide that would enable people to associate linguistic labels with the inclusive concept animal or plant. In fact, many prominent large-scale early civilizations lack unique beginner terms. The Egyptians, for example, while clearly possessing the conceptual category of animal (there is a dedicated classifier for animal labels), did not have an inclusive lexicon term for it (Goldwasser, 2002). Similarly, ancient Chinese also did not have a stable lexical correspondent to the inclusive concept animal. In fact, the historical evolution of the certain inclusive biological terms in ancient Chinese is rather illustrative of the instability of inclusive labels, as will be discussed below.

Rise and Decline of Inclusive Biological Labels: “Chong” (蟲) in Traditional Chinese

Classical Chinese may be viewed as a rare uninterrupted literary tradition lasting more than 3,000 years (Coulmas, 1989). Systematic classification of natural objects occurred in the Chinese language very early: the first comprehensive dictionary, *Erya*, compiled in the fourth to second centuries BC, classified everyday concepts and objects into larger categories. The last seven chapters of *Erya* deal entirely with living organisms and correspond nicely to seven life-forms: grass/herb (*cao*, 草), tree (*mu*, 木), wug (*chong*, 蟲), fish (*yu*, 魚), bird (*niao*, 鳥), wild beast/mammal (*shou*, 獸), and domestic animals (*chu*, 畜), with no mention of any character that has the inclusive meaning of the concept animal. In a slightly later text, *Family Sayings of Confucius*, compiled during the second century BC,¹⁸ the character *chong* is used in an inclusive manner that resembles a unique beginner term:

¹⁸ Kongzi Jiayu 孔子家语 (Family Sayings of Confucius). Zhongguo Wenshi Press. Accessed March 2023 via <https://ctext.org/kongzi-jiayu/zhi-pei>. Scholars debate the authenticity of this text. Many Chinese scholars think the texts were a forgery by Wang Su (AD 195–256) (Kramers, 1950). Whether the texts were really Confucius’ sayings does not matter for our purposes.

There are three hundred and sixty types of *chong* with feather, and *Fenghuang* is their kind par excellence;¹⁹ three hundred and sixty types of *chong* with fur, and *Qilin* is their kind par excellence;²⁰ three hundred and sixty types of *chong* with chitin, and turtle is their kind par excellence; three hundred and sixty types of *chong* with scale, and Dragon is their kind par excellence; three hundred and sixty types of *chong* with bare skin, and human is their kind par excellence. (from *Family Sayings of Confucius*)²¹

From this text it is clear that *chong* is used to represent the entire Animal Kingdom. Even humans are included! This particular classification system is sometimes used in other transmitted texts as well, yet in the same time period some sources use *chong* to refer specifically to wug. In his influential book *Luxuriant Dew of the Spring and Autumn Annals*, the most important Confucius scholar in the early Han Dynasty (Dong Zhongshu, 179–104 BC) uses *chong* in the following way: “poisonous *chong* does not sting; fierce beast does not fight” (毒蟲不螫, 猛獸不搏). Other common phrases in the book, such as “bird, beast, *chong*, and snake” (鳥獸蟲蛇) and “bird, beast, and all kinds of *chong*” (鳥獸昆蟲), also strongly suggest that *chong* simply refers to the wug life-form category.

There are two possibilities here. The first is that the word *chong* is polysemous and refers to both the life-form concept wug and the more inclusive unique beginner concept animal; the second possibility is that there is genuine variation in how individuals associate concepts with the word *chong*. The historical texts itself are insufficient to adjudicate between these possibilities since most people in this period were illiterate and did not leave any written record (Li & Braner, 2012), and as a result we do not know whether they associate *chong* with both concepts in a polysemous manner or they only associate the label with individual conceptual categories. I suspect the second possibility is closer to reality because most occasions in which the inclusive concept animal is invoked are highly philosophical texts such as specifying natural categories or the order of the universe, and ordinary people were unlikely to encounter these texts and as a result unlikely to treat *chong* as polysemous.

Although *chong* may be polysemous in this historical period among some literati, in later periods its animal sense dramatically decreased even in transmitted texts. Using the Chinese Text Project (ctext.org), a comprehensive digital library of pre-modern Chinese texts as a data source,²² I have counted the occurrences of “*chong* with feather” (羽蟲) and *niao* (bird, 鳥) as well as “*chong* with fur” and *shou* (beast, 獸) in its main database (139 texts) by

¹⁹ Fenghuang (鳳凰), a mythological creature in Chinese culture. Frequently translated as “phoenix,” fenghuang only has a superficial resemblance to the Western counterpart.

²⁰ Qilin (麒麟), a mythical hooved chimerical creature known in Chinese and other East Asian cultures.

²¹ Original text: 羽蟲三百有六十, 而凤为之长; 毛蟲三百有六十, 而鳞为之长; 甲蟲三百有六十, 而龟为之长; 鳞蟲三百有六十, 而龙为之长; 倮蟲三百有六十, 而人为之长。

²² The Chinese Text Project comprises an extremely wide range of texts, with thousands of books digitized using optical character recognition. Its original focus was pre-Han philosophical texts, and over time other subjects were added, such as official historical records, medicine, poetry, and fiction.

dynasty and calculated the ratio of the two expressions of the same concept for both pairs in Table 3.

Notice that the relative use of “*chong* with feather” and “*chong* with fur” substantially decreased after the Han dynasty, as indicated by the *chong with feather/niao* and *chong with fur/shou* ratios. In fact, most of the post-Han occurrences of *chong with feather* and *chong with fur* were later authors referencing pre-Han texts. The meaning of the character *chong*, on the other hand, seems to have been largely fixed at wug. For instance, in the highly acclaimed novel *Dream of the Red Chamber* in mid-Qing dynasty, *chong* occurred 24 times, with none of the occurrences referring to the inclusive concept animal. Wang (2012) points out that the shrinkage in meaning of *chong* probably was achieved in the Tang dynasty (AD 618–907) since later authors often needed to provide explicit annotation of *chong* (that it is a general term for bird, beast, and the like) when commenting on earlier texts. In other words, the polysemy of *chong* gradually disappeared, and by the end of the eighth century AD it was almost exclusively associated with the concept wug. Another important takeaway from Table 3 is that the use of *niao* and *shou* vastly outweighs the use of *chong with feather* and *chong with fur*, suggesting that there were significantly fewer opportunities for people to encounter situations in which they can infer the inclusive meaning of *chong*. To summarize, although some early authors clearly used the character *chong* to represent animal, this label failed to maintain the inclusive meaning over time.

The Invention and Spread of *Dongwu* (動物): a Case of Cultural Transmission

It is well known that Mandarin, the official language that people in China use today, is the result of intentional, top-down efforts of phonetic and lexical standardization as well as character simplification since 1949 (Zhou & Sun, 2004). The transition from classical Chinese (文言文) to written vernacular Chinese (白話文), however,

Table 3 Occurrence of the linguistic label *chong* with feather/*niao*, *chong* with fur/beast, and their respective ratios

Dynasty	Date	<i>chong</i> with feather	<i>niao</i> (bird)	<i>chong</i> with feather / <i>niao</i> ratio	<i>chong</i> with fur	<i>shou</i> (beast)	<i>chong</i> with fur / <i>shou</i> ratio
Pre-Han	before 220 CE	24	1900	0.0126	23	1447	0.0159
Wei Jin	266–420 CE	0	325	0	0	160	0
Sui Tang	581–907 CE	3	1114	0.0026	0	776	0
Song Ming	960–1644 CE	10	3274	0.0031	13	2248	0.0058
Qing	1636–1912 CE	4	5952	0.0006	8	1579	0.0051

The data are obtained by searching keywords in the main database of Chinese Text Project and counting their occurrences

was already underway during the Ming (AD 1368–1644) and Qing (AD 1636–1911) dynasties and culminated in the New Culture Movement in early 1900s, when numerous intellectuals engaged in fierce debate regarding the reform and even abandonment of traditional Chinese culture (Rickett & Tse-tsung, 1961; Schwarcz, 1986). The Late Qing period also witnessed a massive amount of knowledge transfer from the West and notably, Japan (Qi, 2004), a country that rapidly modernized after the Meiji restoration (Umetani, 1964).

The modern Chinese word for the inclusive concept of animal, *dongwu* (動物), was in fact coined by Japanese scholars in the process of translating Western scientific work into Japanese shortly after the Meiji restoration in the 1870s,²³ and then imported into China in the early twentieth century, primarily through the introduction of new textbooks in science and technology (Gu, 2009). In Zhang's (2017) analysis of the most influential journal in this period, *New Youth*, *dongwu* appears 15 times from 1915 to 1926, 14 of which are in scientific contexts.²⁴ The use of *dongwu* did exist in classical texts, but its meaning seems to be the literal combination of *dong* (move) and *wu* (thing), which roughly means “thing that moves.” As such, it includes not only living creatures but also individual body parts such as heart²⁵ and leg,²⁶ and even heavenly bodies!²⁷

In the pre-modern period, *dongwu* was sometimes used in a definitive manner to refer to animal. For example, in the same book that uses *dongwu* to refer to heart, leg, and stars in the sky, the same characters are also used in the following context:

[Someone] asks: “*dongwu* have consciousness, yet plants do not have consciousness, why?” The answer: “*dongwu* have blood and vital energy, therefore they have consciousness. Plants, though [we] cannot say they have consciousness, [we] can observe its livelihood.”²⁸ (*A Collection of Conversations of Master Zhu* 朱子語類, AD 1270; Zhu, 1986)

This use of *dongwu* that invokes the inclusive concept of animal, however, is very rare. The only other occurrence in the entire Chinese Texts Project main database (excluding later authors referencing earlier texts) is the following:

In the method of earth to tell the kinds of living things of the Five places: First, mountains and forests, and *dongwu* there are primarily things with fur. . . . Second, rivers and lakes, and *dongwu* there are primarily things with scale. . . . Third, hills, and *dongwu* there are primarily things with feather. . . . Fourth, waterside flatland, and *dongwu* there are primarily things with chitin. . . . Fifth,

²³ During the 1870s, more than fifty works of science and technology were translated from European languages to Japanese (Meade, 2015).

²⁴ The one exception is 政治動物 (political animal).

²⁵ “心本是个動物, 怎教它不动”, *A Collection of Conversations of Master Zhu* (Zhu, 1986).

²⁶ “胙亦是動物, 故止之”, *A Collection of Conversations of Master Zhu* (Zhu, 1986).

²⁷ “日月星辰积气, 皆動物也”, *A Collection of Conversations of Master Zhu* (Zhu, 1986).

²⁸ Original text: 问: “動物有知, 植物无知, 何也?” 曰: “動物有血气, 故能知。植物虽不可言知, 然一般生意亦可默见。”

plains, and its *dongwu* there are primarily things with bare skin.²⁹ (*Zhouli zhushu* (周禮注疏), second century BC; Zheng Xuan et al., 1815)

Notice that this is the same classification system as the one that appeared in *Family Sayings of Confucius* which was compiled at around the same time period, except here *dongwu* is used instead of *chong* to refer to the inclusive concept of animal. This proposed use of *dongwu* did not become the dominant meaning, just as the case with *chong*, until the reintroduction from the Japanese in the early twentieth century. The following historical anecdote may be particularly illustrative of this point: in the year 1864, as the legal interaction between the Qing Empire and the West increased, *Elements of International Law* was translated into Chinese by the American missionary William Martin with the assistance of a number of Chinese officials and legal experts (Tian, 1999). In this important book, *dongwu* was used to translate “movable property.”³⁰ This word choice was not a careless decision or the result of linguistic incompetence; the meaning of texts was carefully explained to Chinese translators who then would discuss and decide the wording to make the sentences sensible to a Chinese audience (Zhang & Zhao, 2009). If *dongwu* had firmly established its meaning as animal, it would not have been chosen to represent a completely different ontological category.

Urban (2010) suggests that the Chinese unique beginner *dongwu* spread into Japanese, Vietnamese, and other Asian languages because of cultural dominance. My analysis here shows that the linguistic label *dongwu* representing the inclusive concept of animal in fact traces its origin to Japanese scholars’ translation of Western scientific work in biology. Like religion, biological theory serves as an important mechanism for maintaining the inclusive meaning of linguistic labels since associating a label with the inclusive concept animal is a relatively weak cognitive factor of attraction compared with associating a label with concepts at the level of folk generics or life-forms, and labels tend to lose their inclusive content over time.

The above historical textual analysis echoes two important points from the previous ethnographic study. First, it highlights the instability of inclusive yet rarely used labels from a temporal perspective. Although *chong* was used by some early authors to denote a rather inclusive category (which included even humans!), such usage did not persist and the lexical meaning of the character experienced a “shrinkage”; second, it provides strong evidence that the inclusive meaning of the term that denotes the concept animal (*dongwu*) in modern Mandarin was crucially facilitated by a powerful cultural institution, the formal educational system.

²⁹ Original texts: “以土会之法辨五地之物生:一曰山林,其動物宜毛物... 二曰川泽,其動物宜鳞物... 三曰丘陵,其動物宜羽物... 四曰坟衍,其動物宜介物... 五曰原隰,其動物宜裸物...”

³⁰ *Dongwu*’s counterpart, *zhiwu* 植物, which in modern Chinese refers to the inclusive concept plant, was used to translate “real property.”

Discussion

Humans possess innate concepts through millions of years of evolution and acquire many others during their lifetime to better adapt to their ecological and social environment (Carey, 2009). The presence of these mental concepts, however, does not guarantee the existence of the corresponding linguistic labels. This is the direct result of how individuals acquire language in natural settings: naive individuals largely rely on their inferential capacities to “guess” the meaning of some linguistic labels, and it is rarely the case that the situations in which words occur definitively determine their meaning. Although we typically assume that individuals’ repeated interactions with their environment lead to label-concept agreement at the community level, such convergence of meaning is not a logical necessity. Cultural attraction theory provides a nice explanatory framework to understand this phenomenon: out of the infinite number of inferential possibilities, biases exist that significantly reduce the number of possibilities and thus greatly facilitate the inferential process. Nonetheless, while word meaning tends to gravitate toward “factors of attraction” (in this case, label-concept matching at the level of entire biological ranks), additional mechanisms are required to ensure the uniformity of meaning in the presence of multiple factors of attraction. Granted, in the equilibrium state one factor of attraction may win out eventually, but there are good reasons to think that real human societies are not in such an equilibrium state (Anzola et al., 2017; Loye & Eisler, 1987), especially with regard to knowledge distribution (Oseledchik et al., 2017).

As mentioned, unique beginner terms have virtually no practical use in the everyday life of people in traditional societies (and of most people in modern societies). Although the concept of animal is likely to emerge very early developmentally (Golinkoff & Halperin, 1983; Ross, 1980), its corresponding label may nonetheless not exist, as shown by extensive ethnographic evidence. In a way, this should not be a surprise given that folk generics is the more salient biological rank. Even in large-scale, complex societies such as China, where thinkers explicitly theorized about the classification of living organisms, the inclusive meaning of unique beginner terms failed to establish. Early authors’ use of unique beginner terms was abandoned by later authors even as China presumably experienced an increase in societal scale and complexity.

I have proposed that cultural institutions such as religion and modern education can serve as powerful mechanisms to help maintain the inclusive meaning of labels that represent plant and animal. In the case of religion, it does so by providing word-learning occasions in the holy texts wherein one can unambiguously infer the inclusive meaning; in education, explicit definitions are often directly taught. This is not to say that religion and education guarantee 100% homogeneous understanding of these terms; in fact, education researchers in the 1980s found that even senior secondary-level students (15–16 years old) in New Zealand fail to count spider, worm, or fish as *animal*, which the researchers point out as being plainly an educational problem (Bell & Barker, 1982). Rather, the presence of these institutions makes a uniform understanding of the inclusive meaning of unique beginner terms more likely.

To recap, the instability of inclusive folk-biological labels has two related consequences. First, individuals are likely to have heterogeneous understanding of these inclusive labels because of their idiosyncratic experience with regard to the situation in which these labels are invoked (as seen in the ethnographic data); second, even if at some point the inclusive usage was established, the semantic inclusivity is likely to get lost owing to a lack of regulatory cultural forces. I highlight the importance of cultural institutions in “regulating” the inclusive meaning of folk-biological labels, in particular how these institutions counter instability by providing unambiguous situations for lexical learning.

Although my field data do not speak directly to the issue of macro-scale language evolution, it has important implications. For one thing, variation in understanding of meaning is necessarily the first step in any semantic change when linguistic authorities are absent. Second, the process of homogenization in meaning cannot be fully understood without considering the role of cultural institutions. Much effort has been devoted to understanding institutions (National Research Council et al., 1988), with some suggesting the spread of certain institutions to be a group-selection type of evolutionary process (Henrich et al., 2012; Norenzayan et al., 2016; Richerson et al., 2016). The present paper thus offers new research possibilities on the study of language evolution, in particular semantic change, by turning to theories of culture.

Conclusions

I have provided ethnographic and historical textual evidence suggesting that the lack of unique beginner terms and the nonuniform understanding of the meaning of inclusive biological labels in small-scale societies may in fact be two puzzles that share the same answer. Unique beginner terms’ salience is low relative to other biological ranks and often relies on cultural institutions such as religion and education to maintain their inclusive meaning.

Acknowledgements I thank Joseph Henrich for his continued support for this project, and Manvir Singh, Ivan Kroupin, Cammie Curtin and two anonymous reviewers for their thoughtful comments and feedback on earlier drafts of this manuscript.

Data Availability All data are available at https://github.com/kevintoy/language_evolution.

Declarations

Competing Interests None.

References

- Andronis, M. A. (2003). Iconization, fractal recursivity, and erasure: Linguistic ideologies and standardization in Quichua-speaking Ecuador. *Proceedings of the Eleventh Annual Symposium about Language and Society—Austin. Texas Linguistics Forum*, 47, 263–269.

- Anzola, D., Barbrook-Johnson, P., Salgado, M., & Gilbert, N. (2017). Sociology and non-equilibrium social science. In J. Johnson, A. Nowak, P. Ormerod, B. Rosewell, & Y. C. Zhang (Eds.), *Non-equilibrium social science and policy* (pp. 59–69). Springer.
- Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, 21(4), 547–569. <https://doi.org/10.1017/S0140525X98001277>
- Atran, S. (1999). The universal primacy of generic species in folkbiological taxonomy: Implications for human biological, cultural, and scientific evolution. In R. Wilson (Ed.), *Species: New interdisciplinary essays* (pp. 231–261). MIT Press.
- Bell, B., & Barker, M. (1982). Towards a scientific concept of “animal.” *Journal of Biological Education*, 16(3), 197–200. <https://doi.org/10.1080/00219266.1982.9654458>
- Berko, J. (1958). The child’s learning of English morphology. *Word*, 14(2–3), 150–177. <https://doi.org/10.1080/00437956.1958.11659661>
- Berlin, B. (1977). Speculations on the growth of ethnobotanical nomenclature. In Ben G. Blount & M. Sanches (Eds.), *Sociocultural dimensions of language change* (pp. 63–101). Academic.
- Berlin, B. (1992). *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies*. Princeton University Press.
- Berlin, B., Breedlove, D. E., & Raven, P. H. (1968). Covert categories and folk taxonomies. *American Anthropologist*, 70(2), 290–299. <https://doi.org/10.1525/aa.1968.70.2.02a00050>
- Berlin, B., Breedlove, D. E., & Raven, P. H. (1973). General principles of classification and nomenclature in folk biology. *American Anthropologist*, 75(1), 214–242. <https://doi.org/10.1525/aa.1973.75.1.02a00140>
- Bloom, P. (2002). *How children learn the meanings of words*. MIT Press.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. University of Chicago Press.
- Bradley, D. (2004). Endangered Central Ngwi languages of northwestern Yunnan. Keynote presentation at the 37th International Conference on Sino-Tibetan Languages and Linguistics, October, Lund.
- Bréal, M. (1904). *Essai de sémantique: science des significations*. Hachette.
- Brown, C. H. (1984). *Language and living things: Uniformities in folk classification and naming*. Rutgers University Press.
- Brown, C. H., Anderson, E. N., Bulmer, R., Drechsel, P., Ellen, R. F., Hays, T. E., Headland, T. N., Howe, L., Hyndman, D. C., Jensen, K. E., Morris, B., & Reason, D. (1985). Mode of subsistence and folk biological taxonomy [and comments and reply]. *Current Anthropology*, 26(1), 43–64. <https://doi.org/10.1086/203224>
- Brown, C. H., Anderson, E. N. Jr., Berlin, B., Boster, J. S., Schadeberg, T. C., & Visser, L. E. (1986). The growth of ethnobiological nomenclature [and comments and reply]. *Current Anthropology*, 27(1), 1–19.
- Buskell, A. (2017). What are cultural attractors? *Biology and Philosophy*, 32(3), 377–394. <https://doi.org/10.1007/s10539-017-9570-6>
- Callanan, M., Cervantes, C., & Loomis, M. (2011). Informal learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 646–655. <https://doi.org/10.1002/wcs.143>
- Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). MIT Press.
- Carey, S. (1985). *Conceptual change in childhood*. MIT Press.
- Carey, S. (2009). *The origin of concepts*. Oxford University Press.
- Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. A. Hirschfeld, & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 169–200). Cambridge University Press.
- Castro, L., Medina, A., & Toro, M. A. (2004). Hominid cultural transmission and the evolution of language. *Biology and Philosophy*, 19(5), 721–737. <https://doi.org/10.1007/s10539-005-5567-7>
- Claidière, N., & Sperber, D. (2007). The role of attraction in cultural evolution. *Journal of Cognition and Culture*, 7(1–2), 89–111. <https://doi.org/10.1163/156853707X171829>
- Clark, E. V. (2017). Semantic categories in acquisition. In Henri Cohen & Claire Lefebvre (Eds.), *Handbook of categorization in cognitive science* (2nd ed., pp. 397–421). Elsevier. <https://doi.org/10.1016/B978-0-08-101107-2.00017-8>
- Coley, J. D., Medin, D. L., & Atran, S. (1997). Does rank have its privilege? Inductive inferences within folkbiological taxonomies. *Cognition*, 64(1), 73–112. [https://doi.org/10.1016/S0010-0277\(97\)00017-6](https://doi.org/10.1016/S0010-0277(97)00017-6)
- Coulmas, F. (1989). *The writing systems of the world*. Basil Blackwell.
- Fodor, J. A. (1975). *The language of thought*. Thomas Y. Crowell.
- Fodor, J. A. (1983). *The modularity of mind*. MIT Press.

- Gasparri, L., & Marconi, D. (2019). Word meaning. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019 edition). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2019/entries/word-meaning/>
- Geraerts, D., Kristiansen, G., & Peirsman, Y. (2010). *Advances in cognitive sociolinguistics*. Mouton de Gruyter.
- Gelman, S. A., & Roberts, S. O. (2017). How language shapes the cultural inheritance of categories. *Proceedings of the National Academy of Sciences*, 114(30), 7900–7907. <https://doi.org/10.1073/pnas.1621073114>
- Goldwasser, O. (2002). *Prophets, lovers and giraffes: Wor(l)d classification in ancient Egypt*. Otto Harrassowitz Verlag.
- Golinkoff, R. M., & Halperin, M. S. (1983). The concept of animal: One infant's view. *Infant Behavior and Development*, 6, 229–233. [https://doi.org/10.1016/S0163-6383\(83\)80030-7](https://doi.org/10.1016/S0163-6383(83)80030-7)
- Gu, J. (2009). 晚清民初日语词汇进入汉语及其启示 (in Chinese) (Borrowing words from Japanese in the late Qing Dynasty and early period of the Republic of China and its apocalypse) Wan Qing Min Chu Ri Yu Ci Hui Jin Ru Han Yu Ji Qi Qi Shi. 深圳大学学报(人文社会科学版). *Journal of Shenzhen University (Humanities and Social Sciences Edition)*, 026(4), 91–97.
- Heintz, C. (2018). Cultural Attraction Theory. In H. Callan (Ed.), *The international encyclopedia of anthropology*. <https://doi.org/10.1002/9781118924396.wbiea2311>
- Henrich, J., & Boyd, R. (2002). On modeling cognition and culture. *Journal of Cognition and Culture*, 2(2), 87–112. <https://doi.org/10.1163/156853702320281836>
- Henrich, J., Boyd, R., & Richerson, P. J. (2012). The puzzle of monogamous marriage. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1589), 657–669. <https://doi.org/10.1098/rstb.2011.0290>
- Hume, D. (2003). *A treatise of human nature*. Dover Publications/Courier Corporation. (Originally published in 1739).
- Hunn, E. (1977). *Tzeltal folk zoology: The classification of discontinuities in nature*. Academic.
- Hunn, E. (1982). The utilitarian factor in folk biological classification. *American Anthropologist*, 84(4), 830–847. <https://doi.org/10.1525/aa.1982.84.4.02a00070>
- Kramers, R. P. (1950). *K'ung Tzu Chia Yü: The School Sayings of Confucius. Introduction, translation of sections 1–10 with critical notes*. E. J. Brill.
- Li, F., & Branner, D. P. (Eds.). (2012). *Writing and literacy in early China: Studies from the Columbia Early China Seminar*. University of Washington Press.
- Liu, D. (2019). The influence of Christianity on Lisu people's cultural development. In *4th International Conference on Contemporary Education, Social Sciences and Humanities (ICCESSH 2019)* (pp. 2046–2050). Atlantis Press. <https://doi.org/10.2991/iccessh-19.2019.438>
- Locke, J. (1847). *An essay concerning human understanding*. Kay & Troutman. (Originally published in 1689).
- Loye, D., & Eisler, R. (1987). Chaos and transformation: Implications of nonequilibrium theory for social science and society. *Behavioral Science*, 32(1), 53–65. <https://doi.org/10.1002/bs.3830320107>
- Macnamara, J. (1982). *Names for things: A study of human learning*. MIT Press.
- Malt, B. C., & Majid, A. (2013). How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4, 583–597. <https://doi.org/10.1002/wcs.1251>
- Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, 16(1), 1–27. [https://doi.org/10.1016/0010-0285\(84\)90002-1](https://doi.org/10.1016/0010-0285(84)90002-1)
- Meade, R. (2015). Translating technology in Japan's Meiji Enlightenment, 1870–1879. *East Asian Science Technology and Society: An International Journal*, 9(3), 253–274. <https://doi.org/10.1215/18752160-3120392>
- National Research Council, Division of Behavioral and Social Sciences and Education, Commission on Behavioral and Social Sciences and Education, Committee on Basic Research in the Behavioral and Social Sciences, Gerstein, D. R., Luce, R. D., Smelser, N. J., et al. (Eds.). (1988). *The behavioral and social sciences: Achievements and opportunities*. National Academies Press.
- Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., & Henrich, J. (2016). The cultural evolution of prosocial religions. *Behavioral and Brain Sciences*, 39, e1. <https://doi.org/10.1017/S0140525X14001356>
- Oseledchik, M., Ivleva, M., & Ivlev, V. (2017). Knowledge as a non-equilibrium dynamic system. *Proceedings of the 2nd International Conference on Contemporary Education, Social Sciences and Humanities (ICCESSH 2017)*. <https://doi.org/10.2991/iccessh-17.2017.1>

- Qi, H. (2004). 中国人日本留学史料补遗--兼谈日本亚细亚大学的中国留学生教育 (in Chinese). *Zhong Guo Ren Ri Ben Liu Xue Shi Liao Bu Yi—Jian Tan Ri Ben Ya Si Ya Da Xue De Zhong Guo Liu Xue Sheng Jiao Yu*. 外国问题研究, *4*, 38–44.
- Quine, W. V. O. (2013). *Word and object*. MIT Press. (Originally published in 1960).
- Richerson, P., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E. K., Naar, N., Newsom, L., Ross, C., Smaldino, P. E., Waring, T. M., & Zefferman, M. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, *39*. <https://doi.org/10.1017/S0140525X1400106X>
- Rickett, W. A., & Tse-Tsung, C. (1961). The May Fourth movement: Intellectual revolution in modern China. *Journal of the American Oriental Society*, *81*(3). <https://doi.org/10.2307/595675>
- Roberts, G., & Sneller, B. (2020). Empirical foundations for an integrated study of language evolution. *Language Dynamics and Change*, *10*(2), 188–229.
- Romero, S. (2012). "They don't get speak our language right": Language standardization, power and migration among the Q'eqchi' Maya. *Journal of Linguistic Anthropology*, *22*(2), E21–E41. <https://doi.org/10.1111/j.1548-1395.2012.01146.x>
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439. [https://doi.org/10.1016/0010-0285\(76\)90013-X](https://doi.org/10.1016/0010-0285(76)90013-X)
- Ross, G. S. (1980). Categorization in 1-to 2-yr-olds. *Developmental Psychology*, *16*(5), 391.
- Schwarcz, V. (1986). *The Chinese Enlightenment: Intellectuals and the legacy of the May Fourth movement of 1919*. University of California Press.
- Scott-Phillips, T., Blancke, S., & Heintz, C. (2018). Four misunderstandings about cultural attraction. *Evolutionary Anthropology*, *27*(4), 162–173. <https://doi.org/10.1002/evan.21716>
- Setoh, P., Wu, D., Baillargeon, R., & Gelman, R. (2013). Young infants have biological expectations about animals. *Proceedings of the National Academy of Sciences*, *110*(40), 15937–15942.
- Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Blackwell.
- Sperber, D. (2012). Cultural attractors. In J. Brockman (Ed.), *This will make you smarter* (pp. 180–183). Harper.
- Tian, T. (1999). 丁韪良与《万国公法》 (in Chinese). *Ding Wei Liang Yu < Wan Guo Gong Fa > 社会科学研究*. *Social Science Research*, *05*, 107–112.
- Umetani, N. (1964). The Meiji restoration. *The Annals of Japanese Political Science Association*, *15*, 99–122. https://doi.org/10.7218/nenpouseijigaku1953.15.0_99
- Urban, M. (2010). Terms for the unique beginner: Cross-linguistic and cross-cultural perspectives. *Journal of Ethnobiology*, *30*(2), 203–230. <https://doi.org/10.2993/0278-0771-30.2.203>
- Vapnarsky, V., Ek, E. U., Medin, D., Lynch, E., Atran, S., & Sousa, P. (2001). Folkbiology doesn't come from folkpsychology: Evidence from Yukatek Maya in cross-cultural perspective. *Journal of Cognition and Culture*, *1*(1), 3–42.
- Wang, H. (2012). “虫”; “蟲” 字初义与意符“虫” 旁类属范畴演变考 (in Chinese). “Chong,” “Chong” *Zi Chu Yi Yu Yi Fu*, “Chong” *Pang Lei Shu Fan Chou Yan Bian Kao*. 陕西师范大学学报: 哲学社会科学版. *Journal of Shaanxi Normal University: Philosophy and Social Sciences Edition*, *41*(6), 163–169.
- Wittgenstein, L. (2009). *Philosophical investigations*. John Wiley & Sons. (Originally published in 1953).
- Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245. <https://doi.org/10.1037/0033-295X.114.2.245>
- Ying, F. T., & Ying, F. T. (2009). (2009). The regional development of Protestant Christianity in China: 1918, 1949 and 2004. *China Review*, 63–97.
- Zhang, L. (2017). 《新青年》(1915–1926) 中的日语借词研究 (in Chinese). “Xin Qing Nian” *Zhong De Ri Yu Jie Ci Yan Jiu*. Doctoral thesis, 北京外国语大学, Beijing Foreign Language University.
- Zhang, L., & Zhao, X. (2009). 从动物, 植物到动产, 不动产—近代法律词汇翻译个案考察 (in Chinese). *Cong Dong Wu, Zhi Wu Dao Dong Chan, Bu Dong Chan—Jin Dai Fa Lv Ci Hui Fan Yi Ge An Kao Cha*, *Journal of Henan University of Economics and Law* 河南财经政法大学学报, *1*.
- Zheng Xuan 鄭玄 (comm.), Jia Gongyan 賈公彥 (subcomm.), & Lu Deming 陸德明 (glosses). (1815). *Zhouli zhushu* 周禮注疏 (*Rites of Zhou*). Shisanjing zhushu 十三經注疏, ed. Ruan Yuan 阮元. In *Sibu beiyao* 四部備要. Accessed March 2023 via <https://ctext.org/rites-of-zhou/di-guan-si-tu>
- Zhou, M., & Sun, H. (Eds.). (2004). *Language policy in the People's Republic of China: Theory and practice since 1949*. Springer Netherlands.
- Zhu, X. (1986). 朱熹, Zhuzi Yulei 朱子語類 (*A Collection of Conversations of Master Zhu*). *Zhonghua shu ju*. Accessed March 2023 via <https://ctext.org/zhuzi-yulei/4>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Ze (Kevin) Hong is a behavioral scientist who studies human behavior and culture from evolutionary and cognitive perspectives (Ed.). He obtained a bachelor's degree in biology from Grinnell College, a master's degree in evolutionary biology from the University of Pennsylvania, and a PhD in human evolutionary biology from Harvard University. He is currently an assistant professor in the Department of Sociology at Zhejiang University and a research associate in the Department of Human Evolutionary Biology at Harvard University. Research-wise, Hong uses methodologically diverse approaches (e.g., theoretical modeling, quantitative analysis of historical data, ethnographic fieldwork) to study human behavior and culture, with special attention paid to information processing at the individual level and information transmission at the population level.