4

Your Mind in Mine: Social Cognition in Grammar

Kolik jazyků znáš, tolikrát jsi člověkem. For each language you know, you are a new person.[1]

(Czech proverb)

So far I have been approaching the structural diversity of languages by examining each part of the system in turn – sounds, meaning, grammar. In this chapter I want to give you a different angle on how much languages can vary, fleshing out the insight of the Czech proverb above by showing why it feels so different to be "inside" different languages.

One of the most insightful discussions about what translation means is by the Spanish philosopher Ortega y Gasset:

> *Cada pueblo calla unas cosas para poder decir otras. Porque todo sería indecible. De aquí la enorme dificultad de la traducción: en ella se trata de decir en un idioma precisamente lo que este idioma tiende a silenciar. Pero, a la vez, se entrevé lo que traducir puede tener de magnífica empresa: la revelación de los secretos mutuos que pueblos y épocas se guardan recíprocamente y tanto contribuyen a su dispersión y hostilidad; en suma, una audaz integración de la Humanidad.*[2]
>
> (Each people leaves some things unsaid in order to be able to say others. Because everything would be unsayable. From this follows the enormous difficulty of translation, which sets out to say in a language precisely what that language tends to remain silent about. But at the same time, it can be seen that translation can be a magnificent enterprise: to reveal the secrets that peoples and times keep from one another, and that contribute so much to their separation and hostility – in sum, an audacious integration of humanity.)[3]

The different grammatical choices made by different languages – in what to say and what to be silent about, in Ortega y Gasset's terms – give very different priorities about what to attend to in the world. We could illustrate this with just about any of the Kantian dimensions of experience – space, time, causality. But I will focus instead on another that, although

less explored, is perhaps even more culturally malleable: how we keep track of our social universe and the psychology of its personae.

Languages differ not so much in what you *can* say as in what you *must* say.[4] From the thousands of things we can attend to in the world around us, each language makes a different selection of what gets front-seat treatment as so-called *grammatical categories*, which speakers and hearers need to keep constant track of. Using another of Ortega y Gasset's insights, each language is "exuberant" in some respects, going into loving detail about particular aspects of reality that you need to attend to and encode in most utterances, and "deficient" in others, allowing you to slack off and pay them no mind unless you feel like it. More than anything else, this is what gives each language its own distinct psychological cast, because to speak it you have to attend constantly to facets of the world that other languages let you ignore. We will pursue this language-and-thought angle more in chapter 8, since to demonstrate it convincingly we will need to bring in psychological experiments as well as linguistic facts. For now, though, I would like to concentrate on developing a preliminary "audacious integration" of what the pooled expertise of the world's grammars has to tell us about one domain of reality: psychosocial cognition.

It is increasingly clear that our ability to construct and participate in a shared mental world, to coordinate our attention and our goals, and to keep track of who knows, feels and wants what, lies at the heart of being human. It is this intense sociality that powered our quantum leap out of the company of all other animal species by enabling us to build that constantly evolving shared world we call culture. This achievement rests on an ability to keep constant tabs on the social and psychological consequences of what happens around us. But, although this skill is universal at a generic level, different grammars bring very different aspects of social cognition to the fore. By integrating what the world's languages are collectively sensitive to, we can come up with a much richer picture of human social cognition than any one language alone would give us.

It helps to start by imagining what a language would be like whose grammar makes NO reference to social context. This grammar would happily enable statements like "monkeys throw coconuts" or "all men are mortal," which imply nothing about the social context around. And it is exactly because of their stark social unanchoredness that sentences of this type crop up at the beginning of logic courses. But it doesn't take long for information about our social world to creep in. Once I say "a monkey threw my coconut," or "the prisoner must die," individual social agents have been drawn in, anchoring the event to the here-and-now of you and me communicating – what is generally known as the "speech act."

The "a" in "a monkey" marks its referent either as not previously known to you, and to be identified by me later ("and in fact, it was that monkey that escaped from the zoo"), or as something whose identity is unimportant or unknowable: "I don't know which monkey, though," I could go on to say. "The" in "the prisoner" shows I am confident you will be able to identify who I am talking about – a confidence that depends on me closely monitoring how far you are following my thoughts and previous statements. "My" in "my coconut" indicates not only that one of the participants happens to be the speaker or writer – me! – but also that I am aware of a particular relationship between that participant and the coconut, perhaps of ownership (I bought it this morning) or perhaps merely of interest (I have been looking at it greedily on the fruitstand). And "must" indicates a relation

between the still-just-imagined event of the prisoner dying, and my wishes and powers to influence what other people around me do, by issuing an order or stipulation. If I shift to a question – "must the prisoner die?" – there is still a relationship of ordering or stipulation between the speech act and the described event of the prisoner dying, but now it is primarily you rather than me who is being linked "deontically" to the event. ("Deontic" modification is framing a statement with desires or moral requirements like "may," "must," or "ought.") Either I am inquiring about something you know but I do not – whether the prisoner's death is necessary – or I am seeking to influence you by asking a rhetorical question and implying that I do not want the event to happen.

Explorations like this, of the meaning of such categories as definiteness ("the" vs. "a"), possession ("my," "your" . . .) and "mood" ("can," "must," "may" etc.), have long been a staple for philosophers and linguists trying to work out how meaning can be represented and how inferences can be drawn. As such they underpin other enterprises like the representation of information in automatic translation systems, or reasoning algorithms in artificial intelligence. They are categories that are central in English and other European languages. But, once we look at other languages, we start to see the elaboration of rather different categories.

Take the phrase "my coconut." If you try to translate this into many Oceanic languages – say Paamese,[5] a language of Vanuatu – you realize that English has not yet given enough information, that its grammar is deficient with regard to the general domain of possessive relations, whereas Paamese is exuberant in the sense of paying attention to much more detailed distinctions. What exactly are you trying to say, a Paamese speaker would insist? My coconut, whose flesh I am about to eat – OK, say *ani aak*. My coconut, whose juice I wish to drink? In that case, say *ani emak*. My coconut, that is growing on my land? In that case, say *ani esak*. My coconut, that I plan to use for some other purpose (perhaps sitting on it)? In that case, say *ani onak*. Between *ani* ("coconut") and the suffix *-k* ("my"), as you can see, we need to insert an element setting out the intended use the "possessor" will put the object to. Devices like these are generally called "possessive classifiers" by linguists, because they classify the type of possession relation. But another way of seeing them is that they signal a mix of socially recognized ownership types and intentions. Indeed, the late Terry Crowley, who wrote a fine grammar of Paamese, argues that the grammar is classifying types of social control, not just possession.[6]

And by Oceanic standards Paamese is still at the kindergarten level when it comes to possessive classification. The New Caledonian language Tinrin[7] distinguishes, for example, between "my (body part)," "my (burnable object)," "my (thing, to plant)," "my (fruit)," "my (meat)," "my (chewable or suckable object, like sugar cane)," "my (cannibalistically eaten human flesh)," and various others. It is impossible to say just "your X," or "my X" without deciding which of these types of possession is involved.

Projection of intentions, in fact, is a key part of our ability for "social intelligence," as any good poker player or military strategist knows, and working out just how much intention-attribution goes on in our primate cousins is a hot topic in tracing the evolution of hominid social reasoning. At the most complex level, this ability allows us to invest just about any sign with a rich attribution of meaning, by enabling us to guess at the communicative intentions of our interlocutor. Say I am in a room with you, and you point

to the window. According to the context, this may variously mean "could you open the window?," "could you close the window?," "isn't it a beautiful window?," "what about we try escaping through the window?," "oh no, what if they come in through the window?," "see, they did decide to put in that tasteless window they were talking about after all," or "look, it's snowing outside." My job as hearer is to work out which of these you mean. Human empathy is well-developed enough that we excel at this sort of mind-reading game, and the philosopher Grice made this ability the corner of his fundamental theory of "implicature," which explains how we are regularly able to mean more than we are able to say. With words, as with window-pointing, we can rely on our interlocutors to top up what we have actually said with additional interpretation based on their informed reasoning about what they think we are intending to convey. This ability has been central in enabling humans to evolve ever more expressive languages by growing new signs – in the sense of constantly developing new words that can say more than we could before.

For now, though, let us think about how intentions are depicted. A common way of doing this in English is to use the preposition "for": "she's going to the cash-machine **for** money"; "he cut the branch **for** a slingshot"; "he's waiting **for** his appointment letter"; "they're searching **for** a unicorn"; "she moved around the ballroom looking **for** a dance partner"; "he planned a surprise party **for** his wife." Suppose you are a cartoonist trying to draw one of the above scenes, or an actor trying to mime them out: how do you convey these intentions to your reader? It's pretty hard, because intentions are not visible, leaving aside a few crude physical indicators of thirst or desire. So you might resort to a thought balloon in your cartoon. But when we watch people doing things, we don't see thought balloons – we rely on detailed knowledge of how they behave, partly rooted in a carefully acquired set of shared cultural routines. The use of the same word *for* in English makes us forget that we are using very different heuristics to work out what the person we are watching is planning to do in each case.

If we translate comparable sentences into Kayardild, we find that every one of these situations needs to be represented by a different case suffix on the word or phrase denoting the goal or intention – see table 4.1. (Case suffixes, remember from the last chapter, are obligatory markers allowing you to work out the role of each object or person talked about in a sentence.) One of the ways you need to reprogram your mind if you learn Kayardild is to pay careful attention to the different ways people go about achieving their goals, and to break down the ways we impute intention into a nuanced set of subtypes.

Close to intention is volition: whether people are consciously in control of the actions they carry out. We humans can attach great importance to volition – our decisions about whether an action was carried out on purpose or not may make the difference between a finding of murder and manslaughter, or, more mundanely, between deciding whether a cough outside the door is simply a passer-by with a cold, or a polite person indicating their presence unobtrusively. But English does not force us to indicate this difference in every utterance: if I say "I coughed" I could be reporting an accidental or a deliberate cough. Of course I could add "on purpose" or "despite my efforts not to" to make this clear, but the point is that the grammar lets you off the hook.

Some languages require their speakers to report every action as volitional or otherwise. Newari, a Nepalese language closely related to Tibetan, is one.[8] Newari verbs take a long

Table 4.1 "For" and intention heuristics in Kayardild

Suffix	*Meaning*	*Example*	*Translation*
-marutha	for the benefit of	*Ngada waaja wangarra ngumbanmarutha.*	I sing a song for you.
-janiija	in order to find, of something that is wanted and must be actively sought	*Niya kalajalaja makujaniija.*	He's going around looking for a woman.
-marra	to use for, of something that can be transformed from something else	*Niya kalatha jari thungali wangalmarr.*	He cut the tree root for a boomerang.
-iiwatha	in order to find, of something that can predictably be found at a given place	*Makuwalada warraja bijurriiwatha.*	The women are going for cockleshells (e.g. to a sandbank known to contain them).
-mariija	for, of something that can only be obtained by waiting	*Makuwalada diija balungka wirrinmariija.*	The women are sitting there in the west waiting for their pension money (which arrives at the post office on a known day).
-kuru	for, of something that is an intention in someone's mind but may not actually exist	*Dangkawalada janijanija Barrindindiwuru.*	The men looked everywhere for Barrindindi (a mythical monster).

final *-ā* (shown by the macron over the *a*) if reporting a volitional action, such as *jī jyā yān**ā*** ("I worked"). But they take a short final *a* if reporting a non-volitional action, such as *jī thul**a*** ("I realized, understood") – since realizing, like remembering, is something over which we have no conscious control. Some verbs, like "meet," can occur with either suffix, depending on whether the action is deliberate or not. Say I meet my friend Manoj. If our meeting was planned, I would say *jī mānaj nāpalān**ā***, with a long final vowel, but if our meeting was by chance, I would say *jī mānaj nāpalāt**a***, with the short vowel and the substitution of *t* for *n* just before it.

Since every verb has to be marked for the volitionality contrast, and determining whether the actions of others are carried out on purpose can keep a judge and jury going for many years, you might wonder how Newari speakers manage to apply this to other people's actions. The answer is that the volitionality contrast only comes into play when recounting your own actions, or when questioning those of your interlocutor, who can vouch for their own directly, as in *chā a:pwa twan-**ā** lā?* ("Did you drink too much (of

your own volition)?"). Elsewhere the *-a* form is used, so a better definition of the contrast would be that *-ā* means "volitional, and knowable as such by introspection" while *-a* means "not certifiable by introspection as volitional."

In fact many languages force their speakers to hold back on what mental and emotional states they can attribute to others. Japanese and Korean, for example, both ration the reporting of "private predicates" attributing inner sensations and feelings like "want," "(feel) cold," or "(feel) lonely" to those who can experience them directly.

English is insensitive to this. It is fine to say both "I want to drink water" and "he wants to drink water," "I am cold" and "he is cold," "I am lonely" and "she is lonely." There is no problem in translating the "I" versions of these into Japanese, e.g. *mizu ga nomitai*, literally "water drink-desirable," for "I want to drink water," *samui desu*, literally "cold is" for "I (feel) cold," and *sabishii desu*, literally "lonely is" for "I feel lonely." But I cannot make comparable assertions about other people, since I can never be 100 percent sure what they are wanting or feeling. Rather, I have to use a more circumspect construction closer in its meaning to English "is acting" or "appears to be," e.g. *kare wa mizu o nomita**gatte**iru* ("he evidently wants to drink"), *kare wa samu**gatte**iru* ("he appears to be cold"), or *kare wa sabishi**soo**da* ("he seems to be lonely"). Korean is similar, and also extends the "privacy" condition to verbs like "like." And it skews the translation of statements about people's presumed future actions through the "presumptive" form *kalkeeyyo*, which, used of myself, means "I will go, am going to go," but used of someone else means "he will presumably go, is sure to go." This recognizes the fact that we can be more certain about our own future actions than those of others.

In the examples so far, the question of evidentiary grounding has been restricted to the difference between what can be known subjectively and what is "external" and evident to everyone. But some languages insist on more careful attention to evidence for *all* statements that are made, specifying whether the speaker knows about it from doing it themselves, seeing it, detecting it by some other sense, from hearsay, from inference, or by other means – typically by a grammatical marker on the verb.

Take Eastern Pomo,[9] for example, now spoken by just a few old people in northern California. To translate English "it burned," you have to choose between four suffixed forms of the verb: *p^{h}a·bék^h-**ink'e*** if you felt the sensation yourself, *p^{h}a·bék-**a*** if you have other direct evidence for it, *p^{h}a·bék-**ine*** if you saw circumstantial evidence and are inferring that it happened, and *p^{h}a·bék^h-**·le*** if you are basing your statement on hearsay. It is possible to translate these back into English versions that have the precision of the Eastern Pomo versions: respectively "I felt it burn me," "I saw it burn," "it must have burned," "they reckon it burned." But the point is that in English we don't *have to* do this – we can get away with being sloppy about the grounds for our statements and just say "it burned" for all four situations. In Eastern Pomo, on the other hand, you *must* specify your source of information for all statements made, so speakers are forced to weigh up their evidence carefully every time they say anything. In fact, a study by Martha Hardman[10] of another language with a well-developed evidential system, Aymara in Bolivia, found that a great deal of effort by children's caregivers goes into teaching them the exact conditions under which it is valid to use the different evidential forms, so as to ensure they are scrupulous and accurate reporters of information.

Well-developed evidential systems are found in many parts of the world – Turkey and the Caucasus, the Himalayas, highland New Guinea, and much of the Americas – and over the last couple of decades linguists have been mapping out the evidential systems across these languages. Gradually they have elaborated a robust cross-linguistic framework or *typology* showing what contrasts are made – e.g. which types of sensory evidence are distinguished? what is considered the most reliable evidence? – and whether one evidential can be stacked up on another. Returning to Eastern Pomo, there are examples of storytellers inflecting a verb for BOTH the non-visual and the hearsay evidentials in a story where the speaker attributes to oral tradition the report of an auditory perception by the old man of someone else walking out.

(1) bà-xa-khí xówaqa-nk'e-e.
then-they.say-he outwards.move-NON.VISUAL.SENSORY-HEARSAY
"Then he started to walk out, it is said (the old man villain, who is blind, heard the hero start to walk out)."[11]

Our cross-linguistic understanding of evidentials was starting to settle into a comfortable form when linguist David Fleck began investigating evidentials in Matses, a Panoan language in Amazonia along the Brazilian–Peruvian border, and discovered a whole new unimagined dimension to evidentiality.[12] Fleck had started out as a zoologist, and Matses people had invited him to conduct research on the animals in the jungle around their village. As he tried to learn the language and discovered he had bitten off more than he could chew without the right analytic teeth, he decided to switch to linguistics. When he told zoologist colleagues about his change, many of them asked him if he thought there would really be enough material for a doctoral thesis – "a primitive language, maybe enough for a Master's." The Peruvian soldiers who gave him a lift out to Matses territory had a different story – "an impossible language, nothing in it makes sense." And the Matses themselves told him: "you're lucky to be working on our language: it's nice and straightforward."

Figure 4.1 A Matses hunter returning with freshly hunted peccary[13] (photo: David Fleck)

The astounding thing about Matses is that it can locate both the reported event and the weighing up of evidence separately in time, with independent yardsticks for each. Say a hunter is returning to his

village from the jungle, and reports that white-lipped peccaries (*shëktenamë* in Matses) passed by a particular location, on the basis of inference from seeing their track. The verb for "pass by" is *kuen*, and the end of the word takes a suffix meaning "it/they," which is *ṣh* or *k* according to conditions we needn't worry about here. Now comes the wild part: depending on how much time elapsed between the event and the detection of the evidence, the speaker chooses the suffix *ak* (a short time period), *nëdak* (a long time period), or *ampik* (a very long time period). And depending on how long passed between the detection of the evidence and the report, the speaker chooses *o* (short time period), *onda* (a long time period), or *denne* (a very long time period). The time-to-detection suffix comes first, then the time-to-report suffix. This gives sentences like:

(2) *shëktenamë kuenakoṣh.* White-lipped peccaries (evidently) passed by.
(Fresh tracks were discovered a short time ago.)

(3) *shëktenamë kuenakondaṣh.* White-lipped peccaries (evidently) passed by.
(Fresh tracks were discovered a long time ago.)

(4) *shëktenamë kuennëdakoṣh.* White-lipped peccaries (evidently) passed by.
(Old tracks were discovered a short time ago.)

(5) *shëktenamë kuenakdennek.* White-lipped peccaries (evidently) passed by.
(Fresh tracks were discovered a very long time ago.)

Like many empirical discoveries, the possibility of a system like this is obvious after the fact: if Matses did not exist, some philosopher of language would have had to invent it. And now that we know about the Matses system, we can go on to make sure that representational logics developed for evidential systems, and cognitive models of social reasoning more generally, do not simply classify evidential judgments by type, but locate them in time as well. But the point is that, to my knowledge, no linguist or philosopher HAD actually postulated such a system. There are many more things in the languages of this earth than have yet been dreamed of in our philosophy.

Reviewing the points we have touched on in our quick tour of what different languages elaborate in the domain of social cognition, we can put forward the following overall model of what any language has to enable its speakers to do (see figure 4.2).

First, as shown on the top left of the figure, a language needs to keep a running file on all social agents – people in one's social universe – and the relations between them (e.g. of kinship, clan membership etc.). Sometimes it has to triangulate between more than one of these relationships at the same time, as in Kunwinjku, where the special system called Kun-derbi[14] locates kin from two perspectives at once. Say grandmother Ann is talking with her granddaughter Valda about Mary, who is Ann's daughter and Valda's mother. Ann would refer to Mary as *al-garrng* ("the one who is my daughter and your mother, you being my daughter's daughter"), while Valda would refer to her as *al-doingu* ("the one who is your daughter and my mother, you being my mother's mother"). Here the relations between three people (a, b, and c) are simultaneously specified, and at the same

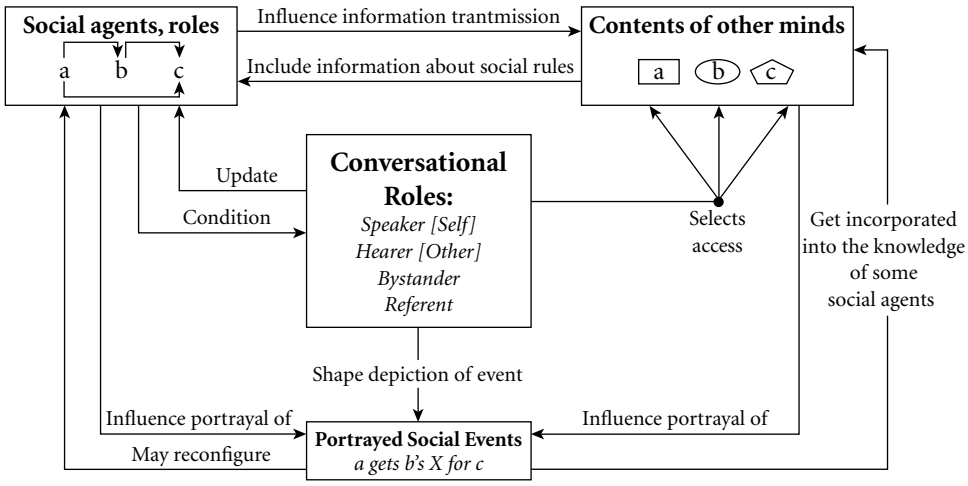

Figure 4.2 A generic model for social cognition in grammar

time linked to the particular conversational roles of speaker, hearer, and referent. A number of Aboriginal languages have systems like this, and their speakers regard the correct use of them, with the ability to take twin perspectives that it requires (plus the encyclopedic knowledge of how everyone in the community is related to everyone else) as the ultimate in courtly refined language.

Passing to the second component in our model, each person in our social universe carries their own set of desires, beliefs, thoughts, and information. Ideally the speaker wants to keep track of all of these so as to predict other people's behavior and interactions. As we saw above, though, many languages are scrupulous in reminding us of the boundaries to how far we can reasonably go in imputing "private" feelings and intentions to others. In the top right of the model, then, are files on what each person, as far as the speaker knows, has in their minds and feelings.

Third, most events that get talked about have some sort of socially relevant component, such as projected intention or change of possession, and the grammar has to allow these components to be depicted in event-descriptions, as shown in the bottom component.

Finally, the act of speech lies at the center of the whole model, as it is speech acts that enable the reliable updating and transmission of such information. A special feature of conversation is its careful training of joint attention, combined with turn-taking between speaker and hearer roles, which permits me as hearer to become privy to your descriptions of your inner feelings. These features enable the intense building of empathy that gives us our best information about where others' minds are at. At the same time, the various types of social relationship impact on conversation and shape the way utterances are framed, something that many languages index grammatically through various types of politeness or respect markers.

To round this chapter off, let us put together some of these themes by seeing how all the socially relevant dimensions of experience we have touched on go into building a

single inflected word in Dalabon, a language that makes you attend to rather different social categories from what English grammar directs you to. (As you can see, Dalabon is a polysynthetic language, like Ubykh, which condenses into a single word what would take a sentence in English.) Here is the word:

(6) *Wekemarnûmolkkûndokan.*
"I'm afraid that the two of them, who are in odd-numbered generations with respect to one another, might go, with consequences for someone else, and without a key person knowing about it; by choosing the form of words I do, I hereby indicate that one of those carrying out the action is a mother-in-law of mine or equivalently respected relative."

To speak Dalabon and use words like this, you need to build them up in the following way. First start with the basic meaning, "go." In Dalabon, this is normally *bon*[15] – but when talking about your mother-in-law or comparable high-respect relatives, you have to use a special polite form of this word, *dokan*. So your first job is to check out the kinship relation you bear to the person you are talking about.

Next you put on the prefix *molkkûn-* ("unbeknownst"), which we discussed in the last chapter – in this context it would denote that the action is being carried out surreptitiously, without letting someone know who should have known about it. So you need to know not just what the two people are doing, but whether they are keeping the right people informed about it. Then you add *marnû-*, which means "for someone's benefit / to someone's disadvantage" or "on someone's behalf" – representing the event as having some effect, positive or negative according to the context, on a third party.

Then you need to specify whether two or more than two people are going, and if there were just two, what kinship relationship holds between them. The word I gave you would be appropriate for talking about a mother–daughter or uncle–nephew team, for example, but to talk about two sisters or a grandmother–granddaughter pair you would replace *ke-* with *barra-*. Finally, you add the "apprehensive" prefix *ke-*, to depict the situation as undesirable. I have translated this here with "I'm afraid that . . . might." But in another context, such as advising a more competent person to make sure the two do not go, or that they let the right people know that they are going, a better translation might be "so that they don't end up" or, in more old-fashioned English, "lest."

I have used this word because it illustrates how far grammars can bring us away from our socially disconnected "monkeys eat bananas" scenario. To plan, utter, or decode the single Dalabon word *wekemarnûmolkkûndokan* we need to bring in all four elements of the model I gave above. Starting at the conversational nexus, I locate the event, in the set of possible worlds, as not corresponding to the here and now, and indicate something of my own attitudes by specifying it as undesirable. Moving to the depicted event, by using the prefix *marnû-* I indicate that this event will have broader social ramifications, bringing benefit or misfortune to others. Now looking at the consequences for the society of minds whose contents we are all engaged in trying to keep track of, by using the prefix *molkkûn-* I depict the event as being unbeknown to someone whom it concerns – perhaps the owner of the clan lands they were on, or the recipient of their planned trophy. As for

the modeling of social relationships, I refer to kinship relations in two places: the prefix *ke-* keeps track of the "odd-numbered-generation" relationship between the hunters, while the use of *bonghmû* instead of the more normal *bon* for the verb root "go" keeps track of my own in-law relationship to someone else present in the conversation.

Intricate as it is, this one-word example only scratches the surface of how languages use their grammars to construct and update their speakers' ever-unfolding dossier of the social universe they move in. Ngalakan, a language next door to Dalabon, has a special "compassion" prefix, which goes on the verb to indicate the speaker's sympathy for someone in the described event. Languages of the Amazon or the New Guinea Highlands would force me to be scrupulous about what grounds I am asserting this on, specifying my source of evidence. Tibetan,[16] or the Ecuadorian language Tsafiki, would have me specifying whether the information being reported is new to me or something I have known for a while, while others like Andoke[17] in Colombia require me to decide whether what I am reporting is something you are likely to have been aware about yourself, or is evident just to me. Although all languages require us to assess the social significance of what we recount and to position the news and the way we present it with respect to ourselves and our conversation partners, the exact demands they make on our minds as we speak, and the way our mental dossiers on our social universe get updated in each conversational move, vary drastically from language to language. With each new grammar we examine, our composite model of how humans are able to reason about the world becomes richer.

To speak Kayardild you need to discriminate many types of intention. To speak Dalabon you have to pay constant attention to the kinship relations between all people in your social world. To speak Japanese or Korean, you must pay close attention to the boundary between what is knowable by introspection and what is knowable by external observation. To speak Newari you need to keep track of volitionality. To speak Eastern Pomo or Matses you must carefully weigh and specify your information source for each statement. Of course English-speakers, as well, can learn to do all these things, and particularly need to do so if they want to function as genealogists (kinship!), psychologists (private predicates!), judges (volitionality!), or well-footnoted academics or journalists (give your sources!) – or, more generally, as empathetic, sensitive, socially switched-on people who are scrupulous about what they say.

How far Kayardild, Dalabon, Newari, Japanese, Korean, Eastern Pomo, or Matses bring this awareness on sooner or more routinely than English does is the sort of Whorfian question that needs a coordination of linguistic and psychological methods of the type to be outlined in part IV, and there has not as yet been significant research in this area.

But what is clear, just from looking at these languages in the detail that we have, is that you cannot speak them without paying constant attention to the particular sets of categories that they force their speakers not to stay silent about. For each of these languages, the speaking cultures that gradually shaped them over millennia must have made these distinctions often enough in past talk by their speakers for them to become installed in their core grammatical apparatus. The occurrence of distinctions like this in a given language thus provides *an existence proof on the learnability and usability of the respective categories* – a proof that not just psychologists, judges, and genealogists, but all normally functioning

members of a speech community, can readily learn these distinctions and incorporate them into more or less permanent attentional scanning. And they also provide *an existence proof on evolvability*: that the words and structures that these grammatical categories derive from can be used often enough for them to evolve into grammatical markers.

A child coming into the world has to have a mind capable of figuring out all of these grammars, and of learning to attend routinely to any of these categories in the course of acquiring their mother tongue. And, as Ortega y Gasset intimated, to map the whole set of human possibilities we need to engage on a bold and vast integration of what the cumulative sensitivities of the world's languages can tell us.

Further reading

Besides Ortega y Gasset's own work, a fine discussion of this problem informed by his position is Becker (1995), with an interesting discussion of Burmese. On the centrality of social cognition to human culture see Goody (1995), Tomasello (1999a, 1999b), and Enfield and Levinson (2006). On evidentiality see Aikhenvald (2004) and the collection of papers in Chafe and Nichols (1986); on the role of inferencing power in building the signs of language see Keller (1994, 1998). Enfield (2002) and Evans (2003b) contain discussions of the mechanisms by which cultural emphases shape grammars over time.