

## How to speak to an alien

---



It is arguably dangerously naïve to suggest that as a species we should send out signals into the universe saying, roughly, “Here we are. Come and decimate our population and enslave the survivors.” A glance at the history of our own species will usually be enough to convince people that when a technologically more advanced society contacts a less advanced one, the outcome is generally not good for the latter, involving as it does, massacres, epidemics and enslavement. The European invasions of the Americas stand out as particular unpleasant examples but there are others elsewhere. Any alien civilisation will, *ex hypothesi*, be more advanced than ours (at least technologically) or they would not be visiting. We may hope that it will be ethically more advanced too but that is not a safe assumption to make.

Shostak, the director of the Center for SETI Research at the Search for Extraterrestrial Intelligence Institute, writing in the New York Times, put it like this:

*Broadcasting is likened to “shouting in the jungle” — not a good idea when you don’t know what’s out there. The British physicist Stephen Hawking alluded to this danger by noting that on Earth, when less advanced societies drew the attention of those more advanced, the consequences for the former were seldom agreeable<sup>1</sup>*

It is to be hoped, therefore, that despite these efforts, none of us will ever have the opportunity to put any of the following ideas into practice face to face (or face to something or other, anyway) with an alien visitor from another star system.

Should we, however, make contact with a distant non-human civilisation, and if we can overcome the 50+-year gap between “Hi, how are you?” and “Fine. You?”, we will need more than the ability to send out mathematical formulae and recordings of Bach to convince any alien interlocutor that we are not to be enslaved and massacred out of hand, at least not for the moment.

We will need language. What follows is a small exercise in what is sometimes called xenolinguistics. If, as we suppose, any alien civilisation which takes the time and trouble to contact (or even visit) us will be much more advanced than we are then it may already have programs in place which can translate our language(s) into theirs without a great deal of difficulty. However, just as it is polite and helps to curry favours from people we encounter on foreign visits to speak at least some of the local language, so it may be considered worthwhile to impress alien visitors with an effort to learn to communicate with them in their language.

In what follows we use terms such as *talk about*, *say* and so on which imply that language has to be orally produced and audibly received. It happens that human language is overwhelmingly oral-aural (with writing a very recent invention, comparatively). There is no *a priori* reason why language has to go by this channel and nothing which follows is inapplicable to other media of transmission. It happens to be the case, however, that vocalisation rather than gesture, grimace, changes in skin patterns and

---

<sup>1</sup> Shostak, S, 2015, *Should We Keep a Low Profile in Space?* New York Times available at <https://www.nytimes.com/2015/03/28/opinion/sunday/messaging-the-stars.html>

colours, organ pipes and so on is a very efficient way of communicating. You can talk in the dark, talk round corners and talk with your hands full. Writing and recording speech gives us the additional ability to talk to people we have never met, after we are dead or who are yet to be born.

### *Assumptions*

What follows is concerned with what an alien language might be like and proceeds from four assumptions of which the first three are:

1. That any race of aliens we encounter will live conceptually in a Newtonian universe of things and not:
  - a. exist at the subatomic level and perceive the universe quantum mechanically
  - b. be non-corporeal having an existence outside of time and space
2. That the law of parsimony will dictate to a large extent how any alien language will function
3. That evolution happens everywhere.

All three assumptions can be questioned, naturally.

There may be intelligent alien entities which live their lives at the subatomic level but it is difficult to imagine how we would interact with such beings and indeed what reason such beings would have for wanting to interact with anything as big as us. It is also imaginable that there exist races of self-aware beings who have no corporeal existence at all and live out their lives (if, indeed their lives ever end) in hyperspace without contact with matter at all. It is equally hard to imagine how we would ever come into contact with such a race at all and if we did, what possible use such a contact could be to either of us.

There may also be parts of the universe where parsimony plays no part at all and where there is no limit to the amount of matter, energy, intelligence, memory and computing power available to the inhabitants. It is not at all easy to imagine such a state of bliss and harder still to imagine why any of its inhabitants would want to go anywhere else or talk to anyone else.

We take it as an assumption that if evolution happens here on earth, that it happens everywhere there is life. We also, incidentally, take it on trust that the laws of physics do not vary across the universe. There may be places where gravity, for example, does not exist but it is hard to make a case for assuming that.

The fourth assumption is, of course, that any alien civilisation we encounter will have something we can recognise as language. The ability to use syntactical language (rather than communicate symbolically in the manner of bees and monkeys) is a defining aspect of being human. It is very hard to imagine a civilisation arising which is capable of contacting another across light years without some means of communicating precise and subtle meanings. If we leave out telepathy, this means some kind of symbolic tokens organised into syntactic units. That is what language is. Even telepathic communication would arguably require **some** use of symbols, of course.

So, proceeding from the four main assumptions, here's what an alien language might be like with some clues for interpreters.

### *Building blocks*

#### *Things*

If we are right in assuming that any species we contact 'out there' lives in a Newtonian universe of things then objects must be referred to somehow. Things exist and language has to have a way of referring to them.

All known languages on earth do this by the attachment of labels and through the invention of nouns to refer to material and immaterial objects. It is not easy to imagine any civilisation advanced enough to journey across light years or to make contact with others light years away could have invented the technology needed without the ability to refer to things. A star ship will, rather obviously, not be referred to as such but it is not rational to assume that its builders will have no way at all of referring to it.

It is probably safe, therefore, to start looking at an alien language from the presumption that there will be a system of object naming.

Of course, even in earthly languages how particular things are labelled is very variable and it is rare to find a neat one-to-one correspondence between languages concerning the area encompassed by any noun. Nevertheless, the systems are close enough for interlingual communication to be usually effective and unambiguous. We would expect any alien language also to have a word (or whatever) to act as the label for a *star* or *liquid water* but what these labels would be is anyone's guess. Such a language might, for example, have no use for a label for *leg* or *vegetable* because the need has never arisen. It may also have labels for things outside human experience, of course, but for that we have invention and neologism. There will also likely be some human-experienced things for which the alien culture has no name and we must hope that their language, too, allows for neologism (it is hard to see how that would not be possible).

It is, however, not a safe assumption that there is a universal need to refer to the non-material world so we cannot expect that all alien languages will be able to refer to *happiness, pain, sorrow, interest, patience, intelligence* and so on but, at least at the outset, this may not be necessary. It may well be the case that an alien species will not experience any of those things and may experience other internal events which have no parallels on earth. We will need to deal with that. The single exception will probably be the concept of *information* for without it no progress in building a vehicle capable of crossing half a galaxy or communicating with people at that remove seems possible.

Now parsimony comes into play. It is conceivable that we will encounter an alien language which can name but not classify so, for example, although all the individual components of the star ship or the teleporter will have names, they will have completely different names in another star ship or teleporter and so on. Such a non-classified naming system would mean that here on earth, for example, each and every occurrence of a motorised vehicle would have a different label attached to it. We could not speak or write of *an automobile* but would have to confine ourselves to talking separately about all of them. That would require something in the order of 1.4 billion different nouns. This would, of course, have to be repeated for all other potentially classifiable entities – *dogs, cats, insects, tables, lights, buildings* and so on – which would make any normal communication impossible because human memory and computing power is not infinite. It would also, incidentally, make the use of an indefinite article or a mass or plural noun impossible because each and every instance of a noun would be *the* item in question. There would be no such concepts as *a car, an insect* or *a tree* and we would not be able, or need, to talk about *cars, insects, trees, water* or *information* because each item or part of each mass would have its own, unique label.

As we noted above, we are working from the parsimony assumption that any alien society will also not have infinite memory and processing capacities so some kind of classification system seems inevitable if object labelling is to be usable.

The question now begging to be answered is how an alien classification system might work. On earth, most languages classify by function so we have words for, for example, *cars, people, houses, goats, ships, substances* and so on which refer to their functions, not their shapes, colours or constituent

materials. We do not, it seems, have languages which classify everything by colour or shape (although some languages do vary in how things are counted by considering characteristics such as animacy, shape and so on). So, instead of saying, e.g., “There’s a car outside the house”, we might have, “There’s a yellow-thing outside the box-thing”, or “There’s a metal-moving-thing outside the brick-not-moving thing”, classifying some things by colour and others by shape, capabilities or material.

There are, in theory at least, few restrictions to the ways in which a language might choose to classify nouns but some classification seems unavoidable if the limits of processing and memory are not to be overstepped. Would-be interpreters are therefore advised to be on the lookout for alternative classification systems.

Additionally, in accordance with the principle of parsimony, it seems that classification by size, potential hazard, edibility, temperature, colour etc. is unlikely to be an effective communication device because we need a classification system that can apply to all entities. Given that there are selective pressures on the development (not evolution in this case) of a language to make it an effective medium of communication, what better classification can a language have than function?

#### *Distinctions between things*

Assuming, for the moment, that any alien language will need a way to label and categorise things, it seems quite likely that it will need a way to distinguish between two things in the same category. Just as human-language speakers can distinguish, for example, between *a glass table* and *a wooden table*, *a big rock* and *a small rock*, *a yellow car* and *a blue car*, *this stone* and *that stone* etc., so any language would seem to need the same ability. How that happens is likely to be unpredictable but would-be interpreters are advised to look out for distinction devices – also known as modifiers.

It is imaginable, that a language could pull this trick off by giving each type of thing a different label so that, for example, *a blue car* is *a thrug* and *a red car* is *a fronch* or whatever and the word (or whatever) for *a big rock* will differ completely from the one signifying *a small rock*. That’s eminently doable but does mean that the number of nouns in the language would increase to a point at which memory and processing power are overstretched. The parsimonious way to get around the problem is to combine describers with the things they describe – syntax, in other words, of which more in a minute.

#### *Events*

Our second category concerns what human languages use to describe things that are or happen – verbs in most languages. We will call these *events*.

If our visitors do, in fact, hail from a setting in which there are things, then as well as naming the things, they will need to have a system of describing events – what things do, where they are and in what relationships they stand to other things.

Human languages have developed to distinguish between doing verbs (*break, drive, fall* etc.), talking and thinking verbs (*speak, explain, imagine, like* etc.) and being verbs (*exist, be like, adjoin, stand, be* etc.) and that seems to be a universal categorisation here on our planet.

There is no reason why any alien civilisation needs to follow a similar system of classification and its language (or languages) could have a number of different ways to classify events. It might, for example, choose to classify events on the basis of the materials which form the subject of those events so, for example, the verb *break* would differ depending on what is doing the breaking or being broken. Equally, it could do so according to the nature of the object of an event so *sawing wood* would be completely differently encoded from *sawing metal*. Some human languages get close to that kind of classification. A huge range of other forms of classification is possible, of course: whether it is done

deliberately or just happens, whether something is dangerous or not, whether something requires labour or not, whether something involves moving things and so on, and on.

Just as with nouns, however, classification seems to be inevitable. One could, of course, imagine a language in which every object would have a different verb to say where it is so, e.g., “The table is in the corner” would have a completely different realisation from “A chair is in the corner”. At this point parsimony strikes, however, and it seems unlikely that memory and processing power would be wasted on having a different verb to describe the same event for each noun in the language. It might have a different verb to describe the same event for each class of object, however, but that would also be non-parsimonious and akin to use a different word for, say, *move*, depending on the class of object (*car, bird, air, cloud, planet* etc.) doing the moving or being moved.

Would-be interpreters are advised, therefore, to be on the lookout for how events are classified.

#### *Distinctions between events*

Just as human languages need a way to distinguish between things given the same name so they also distinguish between events given the same name so, for example, we can say that “The bird is flying” but we can also distinguish types of flying by saying “The bird is flying slowly” or “The bird is flying to the west” and innumerable other characteristics that we wish to attach to the idea of *flying*. The catch-all description of this ability is the attachment of adverbials.

It is not necessary to do it that way and it is perfectly possible to conceive of a language which varies the verb to express different aspects of the event so the word (or whatever) for *flying quickly* would differ from the word meaning *flying westwards* and that would differ from the word meaning *flying quickly upwards* and so on. Again, parsimony would seem to rule that out for the majority of events because of the excessive memory load that would be imposed on users of such a language. Human languages do this but only to a limited extent so we have, for example, distinctions between *run, stroll, gallop* and *walk* (with some encapsulating a notion of speed). However, most languages do not do this for the majority of words used to describe events so we cannot distinguish in a single event word the concepts of *swim quickly* from *swim slowly* or *swim underwater* or between *fall gently, fall rapidly, fall on a hard surface, fall on a soft surface* or *fall sideways*.

What language does, again, is enlist the help of syntax to combine adverbials with verbs. There’s no reason that any alien language would do this in anything like the same way. It could be done, for example, phonologically, by changing the pronunciation of the event word, affixationally by altering the ending, middle or beginnings of event words, by mutation, altering sounds within a word or in many other imaginable and, to us, unimaginable ways. However, the **ability** to do it seems inescapable if a language is to evolve beyond simple event tokens.

Would-be interpreters are advised, therefore, to be on the lookout for how events are modified.

#### *Times and places*

A key distinction between human language and animal communication concerns temporal and spatial displacement. We can, in all human languages, refer to what is here now but we can also refer to what is not here and is not anywhere around just now (although most things have to be somewhere). We can also, incidentally, refer to what is never going to be anywhere at any time because it is a purely imagined concept. We can say, therefore, in any human language:

*There is a unicorn in this garden*

*There was a unicorn in that garden*

*Do you think there might be a unicorn in any of the gardens around here?*

and so on.

Animal communication mechanisms can, demonstrably, communicate notions such as:

*Warning: there's a leopard (monkeys)*

*There's a decent food supply in this direction over there (bees)*

*Keep off my territory (birds and many other animals)*

and so on but they cannot communicate notions such as:

*There was a leopard here last month*

*Have there been any decent food supplies in that direction recently?*

*From next Monday, I will consider this my territory*

and so on.

Even among human languages, how displacement is achieved is bewilderingly variable so we cannot expect to be able to grasp temporal and spatial relationships in an alien language easily. We should, however, expect to find that the language can do these things because visiting us or communicating with us has demonstrated an ability to plan for another time, elsewhere, so there likely has to be a way of referring to it.

Even at a basic here-and-now level, it seems inevitable that some way of encoding these bits of information will be necessary in any developed society. If our jumping-off assumption is that the aliens live in a Newtonian environment is correct then everything material has to be somewhere at some time. Logically, any language used in that environment would need to be able to express where and when something is, happened or behaved.

Would-be interpreters would be well advised to try to see how it is done but shouldn't expect to recognise the method by analogy with any human language.

## Syntax

Chomsky and others of a like mind see language structure as something inherent and encoded to some extent in our genes. The concept of a Universal Grammar suggests that the basic building blocks of syntax that we have referred to above exist in every human language. The view has been challenged but maintains a good deal of academic currency.

It turns out, however, that Universal Grammar is not quite as universal as it sounds and the holders of this theory go on to suggest that because language which has evolved on earth is unique to the planet, we will never be able to unravel the syntax of an alien tongue and we will have to approach it as we would a problem in chemistry or particle physics, *ab initio* and without the prop of our inherited language acquisition devices (whatever they might be).

Chomsky himself averred (1983):

*The same structures that make it possible to learn a human language make it impossible for us to learn a language that violates the principles of universal grammar. If a Martian landed from outer space and spoke a language that violated universal grammar, we simply would not be able to learn that language the way that we learn a human language<sup>2</sup>*

However, for language to function at all it has to combine symbols with syntax and what we have outlined above does not seem, on the face of it, to violate any principles. It is unlikely that we could learn an alien language **in the same way** that we learned our first language but that does not mean that we would be **unable** to learn it especially if we are forearmed with knowledge of what it is capable of doing and what it must be able to do.

---

<sup>2</sup> Chomsky, N, 1983, Interview with John Gliedman, Omni, 6:11, November 1983, available at [https://chomsky.info/198311\\_/](https://chomsky.info/198311_/)

Animal communication sounds do not have syntax but they do have representations as we saw above. To form more than a way of communicating a single instance of an event (such as a claim for territory, a warning or an invitation to set up home) we need to combine symbols with syntax. That way, we can move from just shouting “Leopard!” to saying “This leopard is harmless” or “Don’t worry. That leopard has already eaten” or “This is actually my territory but I’ll make an exception for you because you seem rather attractive”.

One theory of the evolution of language (not languages) on earth is that it began with a simple symbolic system (*danger, fire, food, threat, warning* etc.) encoded in sound or by gesture and evolved syntax as the demands of a social animal with a big brain gave its acquisition a selective advantage. The obvious selective advantage of the ability to use an increasingly sophisticated syntactical system to convey ever more subtle meanings needs little underlining. It may be helpful to be able to say, “That way food.” but to be able to say, “There is a carcass to be scavenged one hour’s walk in this direction but we must hurry before others get there.” would confer a competitive, survival and reproductive advantage on any animal which acquired the knack. If a genetic base for the cognitive processing of syntax does exist, then greater longevity and concomitantly greater reproductive success will lead inevitably to the ability being inherited selectively. In other words, variation will be rewarded and disproportionately inherited if it leads to greater survival and reproductive chances.

If we are right in supposing that evolution happens everywhere there is life and happens for the same reasons, then it would be odd if any alien visitor’s language did not have a logical and penetrable syntactical system which allowed for displacement, modality, contingency and so on. It would be even odder if such visitors had not developed language at all.

So, although the syntactical system of an alien language is likely to be superficially impenetrable, the functional categories which it encodes are not. Briefly and incompletely, any acceptably advanced language needs to be able to:

1. Name things and distinguish things from each other by labelling and modification
2. Name events and distinguish events by their nature (behaviours, states, thoughts etc.)
3. Locate events and things in time (*when, after what, before what* etc.)
4. Set things in spatial parameters saying how they are related to each other (*where, near what, how far* etc.)
5. Allow for the expression of displacement (*not here, there, then, not now, me, you, they, her* etc.)
6. (Probably) allow for the expression of degrees of certainty and obligation – modality, in other terms (*could be dangerous, won’t be safe, Don’t play with that alligator* etc.)

There are many other clever tricks that language allows us to perform such as setting up and maintaining social relationships, passing on cultural knowledge, persuading others to conform to our wishes, organising resources and labour and much, much more.

None of these functions need, however, trouble us too much at the outset because they are not needed until we have at least figured out the syntactical and symbolic bases of the language we encounter.

If we can communicate effectively about things, events, positions and times, all else will follow.