

# LANGUAGE AS HEADACHE

The Linguistics Olympiad is all about getting to grips with the way that language works. **Dick Hudson** explains how to think like a linguist.

**D**o you know how Armenian writing works? Or how Ulwa speakers know which form of the possessive to use? Or, for that matter, how we English speakers draw conclusions from sentences containing words such as *some* and *every*?

Even if you don't already know how Armenian writing works, maybe you could crack the code by looking at a small sample? An interesting challenge – and especially if we told you that it took some brainpower to work out the answer? And that you might get a brainpower Gold if you managed it (and also a bunch of other similar brainteasers)? And that you might learn more about language in general along the way? Welcome to the Linguistics Olympiad!

You may have heard of Olympiads in subjects such as Mathematics and Physics. Well, this is just like these well-established competitions for school pupils, except that all the problems focus on the patterns you can find in language. If

you already find language fascinating, the Linguistics Olympiad is certainly for you; but if you haven't given language that much thought, a few problems from the Linguistics Olympiad will open your eyes to the frustrating beauty of language structure.

Why 'head-ache'? Just try one of the problems that's slightly out of your comfort zone, and you'll see.

For instance, here's the Armenian alphabet:

Ա Բ Գ Դ Ե Ձ  
 Է Ն Թ Ժ Ի Լ  
 Խ Ծ Կ Ը Ղ Պ  
 Ճ Ա յ Ն Շ Ո Չ  
 Պ Ձ Ն Ա Վ Տ  
 Ի Ծ Ի Փ Բ Օ Ֆ



No idea how it works? Well, you do have some idea – we called it an alphabet, so you assume that each letter corresponds to a single vowel or consonant; and that’s the best way to approach the LO problems, by forming working hypotheses. Your next one might be that there is a regular pattern of correspondences between letters and sounds; and if you’re lucky, the correspondences will be a good deal more regular than in English!

Now let’s suppose we show you a photograph of a metro map for Yerevan (the capital of Armenia). That doesn’t in itself throw a great deal of light on the alphabet, but it does suggest more hypotheses. First, words seem to be separated by spaces, just as in English (but not in every writing system). Second, there doesn’t seem to be a contrast between capital and small letters, because what are obviously names seem to have letters at either end that are also found in the middle. And thirdly, we can be sure that, like English, it’s written from left to right – just look at the margins in the three lines at the foot of the photo.

Even after this much detective work you still don’t know how the individual letters translate into sounds. But suppose we also gave you this little story about a girl called Millie:

*Millie has got lost in Yerevan, the capital of Armenia. She is at the station named Shengavit but her friends are waiting for her at the station named Barekamoutyun. Some other stations are called Gortsaranain, Zoravar Andranik, Charbakh and Garegin Njdehi Hraparak. Your answers, based on the metro map on the next page, can help Millie meet up with her friends. Now you’re in business – and the business is serious, because Millie depends on your help.*

The way forward, here and in all the LO problems, is obviously to crack the code. Once you’ve done that, the questions mentioned in the story are easily answered; you can check your answers against the ones on our website (see the list of links at the foot of this article). And, of course, on the way to the answers you’ve actually learned a bit about Armenian.

Ulwa possessives are a very different matter. With Armenian you were faced with a relatively straightforward job of matching Armenian letters with their English equivalents, but in Ulwa you’re right in the middle of a rather exotic grammar. We start with rather important background information:

Ulwa is a language spoken in Nicaragua. It contains quite a few loanwords from English, which is spoken in the Bluefields area of the country ... Ulwa distinguishes between singular and plural “you”, and also distinguishes between inclusive “we” (we including you) and exclusive “we” (we not including you).’

These tips are important, so don’t forget them.

We now look at the basic data in the table opposite, consisting of 27 Ulwa words together with their English translations, and of course the challenge is to work out how Ulwa works. The first thing to notice is that one Ulwa word may correspond to several English words; so the challenge is to break the Ulwa words down into the parts that correspond to individual English words.

For example, *kapakka* means ‘his/her manner’, so one bit must mean ‘manner’, leaving the rest to carry the meaning ‘his/her’. We’re also told that *kapak* means ‘manner’, so we can be sure that *-ka* means ‘his/her’. So far so good. This analysis is confirmed by *kiiika*, meaning ‘his/her stone’; so presumably *kii* means ‘stone’. Moreover it suggests a division of *gaadni*, ‘our (inclusive) god’ into *gaad-ni* (where we recognise the English loan-word *gaad* for *god*). Presumably the Ulwa for ‘our (inclusive)’ is *-ni*.

Unfortunately, life is never that simple in a Linguistics



## Ulwa words

<b>arakbus</b>	gun
<b>askana</b>	his/her clothes
<b>bilamkana</b>	their fish
<b>bilammana</b>	your (plural) fish
<b>diimuih</b>	snake
<b>diikanamuih</b>	their snake
<b>diimamuih</b>	your (singular) snake
<b>gaadni</b>	our (inclusive) god
<b>iibin</b>	heaven
<b>kahma</b>	iguana
<b>kapak</b>	manner
<b>kapakka</b>	his/her manner
<b>karaskanamak</b>	their knee
<b>kiika</b>	his/her stone
<b>kululuk</b>	woodpecker
<b>liima</b>	lemon
<b>mistu</b>	cat
<b>sapaaka</b>	his/her forehead
<b>sikbilh</b>	horsefly
<b>siknibilh</b>	our (inclusive) horsefly
<b>suumanalu</b>	your (plural) dog
<b>paunimak</b>	our (inclusive) tomato
<b>taikinatai</b>	our (exclusive) grey squirrel
<b>taim</b>	time
<b>uumamak</b>	your (singular) window
<b>waikinaku</b>	our (exclusive) moon
<b>wasakanala</b>	their possum

Olympiad. The second Ulwa word is *askana*, meaning ‘his/her clothes’. Where is the expected *-ka*? It is there, but not where we expect it to be, stuck onto the end of the stem: *asna-ka*. Instead, it is right in the middle of the word. The same applies to *-ni*, which we also find in the middle of *siknibilh*, ‘our (inclusive) horsefly’ and *paunimak*, ‘our (inclusive) tomato’. Why?

In other words, what is the rule of Ulwa grammar that puts *-ka* or *-ni* on the end of some words, and in the middle of others? (Technically, these ‘affixes’ are either ‘suffixes’ or ‘infixes’). How does this bit of Ulwa work? This is a question that faces every Ulwa child, so what we are really asking is how this bit of the Ulwa mind works – what rule do Ulwa children discover for themselves?

We can only arrive at the answer by carefully studying the available data. This isn’t just how Linguistic Olympians work, but how Ulwa children and, indeed, children everywhere, learn about their own language.

But what part of the data should we be looking at? Is it a matter of meaning? Try it: our examples so far go like this:

- suffixes: manner, stone, god
- infixes: clothes, horsefly, tomato

Nothing obvious seems to emerge, such as male/female or animate/inanimate.

A much more promising avenue to explore is the shape of the word, when defined in terms of vowels, consonants and syllables – in other words, in terms of phonology, the way sounds are combined. Phonology is an important part of a language’s structure, and although it varies enormously from language to language, most languages organise their phonology round syllables consisting of vowels and consonants. Representing words as a string of C’s and V’s, with brackets to show syllables, is a good way to get beyond our very misleading spelling; so *through* is [CCV] and *Wednesday* is [CVCC] [CV].

Once you start looking in the right place – in this case, in phonology – you start to make progress. Words like *kapak* take

suffixes (e.g. *kapak-ka*, our first example), while those like *asna* take infixes (giving *as-ka-na*). Without the suffix or infix, the first is [CV][CVC] while the second is [VC][CV]. You notice that the first syllables are different, and if you pursue this difference you find that it leads eventually to a rule for deciding whether a possessive affix should be a suffix or an infix. Again, we won’t spoil your fun by giving the answer away, but if you want to check it you’ll find it on the website.

Before we leave Ulwa, it’s worth reviewing what it teaches us. This is important because the Linguistics Olympiad isn’t just fun (though it is that); its more serious aim is to reinforce what schools already teach about language. This ‘knowledge about language’ includes knowledge about language structure, so exploring complex linguistic patterns develops a deeper appreciation of linguistic.

The main point of the Ulwa exercise is what it shows about how grammars get written: by discovering patterns, making the best possible sense of them, and reporting the outcome – just as in a science experiment, where data + analysis = result. Anything you read about grammar has been discovered by someone at some point in the last four thousand years. This includes all the categories we take for granted in talking about language, from ‘consonant’ to ‘suffix’ and ‘possessive’. Language is patterned behaviour, and each society develops rules to control its patterns. Discovering these patterns is what linguists do, so what you’re doing by working out the rules for Ulwa possessives (or Armenian letters) really is a taste of linguistics, the science of language.

One other moral that is worth drawing from our encounter with Ulwa is the enormous variety to be found among languages. Ulwa has infixes, but English doesn't (unless you include the rather rude *abso-bloody-lutely*); Ulwa has different forms for inclusive and exclusive 'we', but English doesn't; Ulwa uses affixes for possession, but English ... well, what about possession in English? We use separate words like *my* and *our*, but we use the 'apostrophe s' for marking possession, as in *John's hat*. The -'s looks like an affix, but unlike Ulwa, it's attached to the possessor, not to the thing possessed. If you look hard enough, you find similarities as well as differences, so every new language that you look at throws some light on the ones you know already. Which, of course, is one of the reasons why linguistics is such fun.

What about English? Linguistics Olympiad problems like these help you to look at your own language in new ways. For our part, we think English is the most interesting language of all – but then, we would, wouldn't we? But seriously, we would expect you to find your own language more interesting than any other, for the simple reason that any analysis throws light on things you already know. And since language dominates not only communication but also general thinking, the light illuminates the very workings of your mind – how you think.

Take the English 'quantifiers' such as *some* and *every*. Quantifiers have profound effects on the logic of what we say. Let's start with a sentence containing no quantifier:

**(1) This student can read Russian.**

If sentence (1) is true, we can be

absolutely certain that (2) is true as well.

**(2) This student can read.**

This is basic logic, and is described in terms of a special relation called 'entailment': sentence (1) entails sentence (2).

Logicians call the relation between (1) and (2) 'upward entailment', building on the metaphor behind any classification tree in which more general notions are located above their sub-categories. For example, the notion 'reading' is higher than 'reading Russian' because reading Russian is a particular example of reading. Upward entailment means that a sentence about reading Russian entails one about simply reading – i.e. lower entails higher.

Now let's replace *this* by the quantifier *every*.

**(3) Every student can read Russian.**

**(4) Every student can read.**

Does (3) entail (4)? Clearly it does, so *every* allows the upward entailment in which reading Russian entails reading. It looks as though *every* behaves like *this*, but there is also an important difference: (3) entails (1), but (1) does not entail (3). If every student can do something, then so can this student; but if this student can do something, it does not follow logically that the same is true of every student. This is an example of 'downward entailment', in which the higher (every student) entails the lower (this student). In short, *every* allows upward entailment on the right (around reading and reading Russian), but downward entailment on the left (around every student and this student).

But the quantifier *no* is different:

**(5) No student can read Russian.**

**(6) No student can read.**

In this case, there is downward entailment on the

right: being unable to read entails being unable to read Russian. And on the left, around no student and this student, we notice that (5) contradicts (1) – if no student can read Russian, then we can be sure that this student cannot read it. This is a special type of downward entailment, just like every except that the higher excludes the lower.

This logical analysis is actually dressed up in the Linguistics Olympiad as a little story about Professor Monotone and his "Astounding Linguistic Knowledge Engine for Making Inferences" (ALKEMI) for drawing entailments in a mechanical fashion. Your task is to find the fault in the code driving ALKEMI which leads to mistakes in its output. Once again, if you're interested in pursuing the example you'll find it (complete with answers) on our website.

Just three examples from the hundreds of problems in the UK and around the world, but they help to illustrate the broad coverage of the Linguistics Olympiad, ranging from exotic languages to English, and from writing, through grammatical forms, to logic. The examples show that the Olympiad tests neither your existing knowledge of the languages concerned, nor your knowledge of linguistics. The focus is always on an essential linguistic ability: to spot and analyse patterns in the data provided; so it favours those who can think analytically; and not surprisingly, some of our best competitors are also mathematicians. But of course the Linguistics Olympiad is all about language, so it provides a wonderful opportunity to engage actively with language structure.

So how does the Olympiad work? The first Linguistics Olympiad took place in Moscow in 1965, and for most of the years since then it has existed only as independent national competitions. However, an International Linguistics Olympiad was launched in 2003, and the number of national Olympiads has increased from six to more than thirty. When we mentioned 'our website' above, we were referring to the webpages of the UK Linguistics Olympiad (UKLO), which is a relative newcomer (dating from 2010). This is what we describe below, but wherever you are in the world, you're never too far from one of the Olympiads.

In schools, UKLO has triggered a great deal of enthusiasm. In 2012, it involved 2,000 pupils from 106 schools, with another 200 schools registered as interested in participating. The numbers have roughly doubled each year since we started, and look like doubling again in 2013. Among our 'regular' schools, we know that a number are now running weekly 'linguistics clubs' where enthusiastic teachers and students of foreign languages and English language, and even students who aren't otherwise studying language, work through past problems.

Academic linguists have shown similar amounts of enthusiasm, which allows us to run the project mainly on the basis of volunteer labour by academics – a large committee, plus an even larger register of volunteer markers. The marking operation has created a new phenomenon (and word): the 'markathon', which is a meeting for volunteer students and staff where the business is to mark a few hundred scripts. A typical markathon is fuelled by

departmentally funded pizzas or chocolates, which helps to maintain student enthusiasm.

One of the special features of the UK's Olympiad, in contrast with most others, is that we offer it at three different levels of difficulty: Foundation, Intermediate and Advanced. The questions overlap across levels, so competitors get a taste of the easier questions in the next level up, but in principle they aim at different age-groups, starting as young as age 11.

The three levels have very different aims and rules. The overall aim of the whole competition is to give children some experience of doing linguistics and of thinking about language, so the more children take part, the better; and the more times each child takes part, the better. Our aim, therefore, is to start young. If we can hook the eleven-year olds, they will have the opportunity to take part every year until they leave school. From that point of view, then, Foundation level is the crucial one; but at this level, there is no national competition so we leave schools to organise the competition in whatever way they think pupils will enjoy most. In many schools, pupils even work together in teams. Nor do we need to ensure neutrality of marking, so we leave teachers to mark their own scripts (with the help of a detailed marking scheme which we supply). Our only role at Foundation and Intermediate levels, apart from providing the test papers and marking scheme, is to provide participation certificates for schools to print to suit their needs.

The Advanced competition, in contrast, is a national competition, so competitors take it under strict exam conditions, and all the scripts are marked by

## Find out more

### Online

The website of the International Linguistics Olympiad [www.ioling.org](http://www.ioling.org)

The website of the UK Linguistics Olympiad [www.uklo.org](http://www.uklo.org)

Try out some past UKLO problems (answers included) ... [www.uklo.org](http://www.uklo.org)

... as well as many, many more past problems from Linguistics Olympiads in other countries [www.ioling.org](http://www.ioling.org)

the volunteers mentioned above. The best 30% are awarded Gold, Silver or Bronze certificates, while the ultimate prize is an invitation to take part in the second round. This is a residential weekend hosted by some university where the top 16 competitors from the first round work with three experienced tutors before taking another, even harder, test to select the team of four to represent the UK at the next International Linguistics Olympiad. In countries as different as Sweden, Poland, the USA and Slovenia, our teams have matched their skills against the best brains from thirty other countries – and collected a respectable collection of medals.

Now it's over to you. If you are at school or college, either sign up for your school's Linguistics Olympiad or persuade a teacher to start one. If you have or know school – or college-age children, encourage them to get involved. And if you're a teacher, join the party! Even if you're none of the above, but are interested in language in general or language puzzles in particular, visit the UKLO site, where you'll find links to many, many more problems. ¶

---

**Dick Hudson** is Emeritus Professor of Linguistics at University College London and the Chair of the UK Linguistics Olympiad committee.

---