

In Praise of Fluffy Bunnies

Copyright © 2012, Richard Forsyth.

Background

Reading John Lanchester's *Whoops!*, an entertaining account of how highly paid hotshot traders in a number of prestigious financial institutions brought the world to the brink of economic collapse, I was struck by the following sentence:

"In an ideal world, one populated by vegetarians, Esperanto speakers and fluffy bunny wabbits, derivatives would be used for one thing only: reducing levels of risk." (Lanchester, 2010: 37).

What struck me about this throwaway remark, apart from the obvious implication that derivatives were actually used to magnify risk rather than reducing it (doubtless by carnivores ignorant of Esperanto), was its presumption that right-thinking readers would take it for granted that Esperanto symbolizes well-meaning futility -- thus highlighting the author's status as a tough-minded realist.

This is just one illustration that disdain for Esperanto in particular, and auxiliary languages in general, pervades intellectual circles in Britain today, as in many other countries. And if you dare to raise the subject of constructed international languages with a professional translator or interpreter be prepared not just for disdain but outright hostility.

Of course professional interpreters are among the most linguistically gifted people on the planet, and can't see why the rest of us shouldn't become fluent in half a dozen natural languages in our spare time. (Not to mention the fact that a widespread adoption of Esperanto, or one of its competitors, would have a seriously negative impact on their opportunities for gainful employment.)

Thus Esperanto has become a symbol of lost causes, to be dismissed out of hand by practical folk.

Yet those risk-junkies busily trading complex derivatives who brought us to the brink of ruin also thought of themselves as supremely practical hard-headed folk. It turned out that they were in the grip of a collective delusion whose effects have impoverished us all. Perhaps they have something to learn from vegetarians and Esperanto speakers.

In the world of supposedly practical folk today, during an intercontinental recession, the European Union spends vast sums of money each year on translating thousands of tonnes of documents into 23 different official languages. The demand for simultaneous interpreters in Brussels, Luxembourg, Strasbourg and at the UN consistently outstrips supply. Meanwhile in the UK, cohort after cohort of schoolchildren emerge from secondary education unable to understand any language

other than their own, often after years of instruction in French, German or Spanish.

"Never mind," retort the anglophone triumphalists, "English is the international language these days."

If you really believe that English is an adequate lingua franca for Europe, let alone the world, try working in a multi-national research project. I spent 2 years as the only native English speaker in an EU project, with English as its official working language, and have been scarred by the experience.

At first glance, this would seem to represent a triumph for the language of Shakespeare and Churchill: our native tongue has conquered the world! Sitting in a meeting, listening to colleagues conversing in Euro-globish heavily laden with mispronounced English jargon, trying to understand and make one's self understood, one starts to realize that this is not the triumph of English after all. It seems more like a devious kind of linguistic ju-jitsu, in which the world takes its revenge for being forced to accommodate monoglot English-speakers by twisting their language into a barbarous dialect which they find awkward and unfamiliar.

Admittedly, English began as a creole, the offspring of a shotgun marriage between Anglo-Saxon and Norman French, but it has come a long way since then, and I personally am very fond of it. The anglicized pidgin that passes for English as an international language isn't the language I love, and it isn't a very effective medium of international communication either.

As it happens, the most eloquent exponent of English as a means of communication that I have ever heard was a Hungarian. But most of us have neither the talent nor the dedication to reach such a height in our mother tongue, still less in a foreign language. We do, however, have sufficient ability to achieve communicative competence in Esperanto within three months; and when we employ it we'll be communicating with others in the same position as ourselves, i.e. second-language users. There won't be the fertile soil for misunderstanding that exists when a native speaker instinctively exploits the quirks of the language or a nonnative speaker makes a small slip of syntax with serious consequences.

Why then does Esperanto remain a fringe cult? Why doesn't the EU insist that all children in Europe spend even a single term learning Esperanto?

Part of the answer must be that, once you accept the idea of a constructed language, there is always the seductive possibility of doing better. At certain points during a course on Esperanto you will come across a construction (such as using the so-called accusative after a preposition to indicate motion) that makes you ask: why did Zamenhof do it that way -- surely that wasn't a good idea? If I want to learn Chinese, I may be daunted by the tonal system, or the thousands of unfamiliar characters, but I have to accept them: that's the way it is. But with an artificial language I'm tempted to think "that should be changed" whenever I come across a difficult or unappealing aspect.

Esperanto was in several respects superior to Volapuk, and the Idists think than Ido is better in many respects than Esperanto. Not everyone agrees. Jespersen -- no mere

dabbler, he -- believed that Novial was better than either.

So it goes on. Hundreds, perhaps thousands, of artificial languages have been proposed in the past couple of centuries. Most never get used in action. In fact, the second most widely used artificial language, after Esperanto, is probably Klingon, which was deliberately designed to sound harsh and be hard to learn!

Only Esperanto, for all its perceived imperfections, has ever sustained a community of users numbering more than a few thousand for more than a few decades. Other international language projects, apparently more elegant in concept (e.g. Interglossa, Lingua Franca Nova), have remained on the drawing board. A list of those that have attracted at least some serious attention is given in the Appendix to this essay.

Thus, early in the 21st century, we arrive at a situation where Esperanto stands as a proof of concept, but has failed to take off. In spelling it approaches the ideal of one character for one phoneme more closely than almost any natural language, consequently it is easy to pronounce from the page. Its grammar is far more regular than that of most natural languages, consequently it can be mastered in a month. Its vocabulary contains a large number of roots found in the major European languages, consequently it doesn't impose a forbidding memory load on adult learners -- provided that their first language is Indo-European. Above all, it has demonstrated repeatedly that international meetings can proceed smoothly without banks of interpreters sitting in cubicles and wires leading into everyone's ears.

Nevertheless it is generally viewed as merely a hobby for cranks. Linguists sneer at it. EU policy-makers would rather pour rivers of taxpayers' money into translation agencies and an endless stream of machine-translation projects that never quite achieve their desired objectives than attempt to introduce Esperanto into the workings of the EU.

Personally, I believe this situation is highly unsatisfactory. I am motivated to attempt to do something about it for two primary reasons:

1. In today's globalized civilization, the need for a common international medium of communication is more urgent than ever before;
2. The strain placed on English in its role as *de facto* international language is turning it into a monstrosity.

Therefore I intend part of my website to play host to yet another effort to devise a constructed auxiliary language for international communication. I plan to kick off the process and with luck enlist some support.

Why should such a quixotic enterprise succeed, when hundreds before it have failed? Well, it might not; but there is one advantage that neither Zamenhof nor any of the early pioneers enjoyed, and which none of the more recent interlinguists seem to have exploited -- the computer.

Take my Word for it!

An international language needs (1) a simple orthography, (2) a regular grammar, and

(3) an easily learned vocabulary. Typical interlanguage projects tend to emphasize the first two points but leave the third in the background. Yet choice of lexical units is the most important of the three. It is normal for proponents of an auxiliary language to claim that its vocabulary is 'international' in some sense but the foundation for this claim is almost invariably subjective.

Zamenhof's approach to Esperanto vocabulary-building can be described as 'eclectic'. It has been said that Esperanto sounds like a Czech speaking Italian. He selected a motley collection of roots from the Germanic, Romance and Slavic languages of Europe. The effect is not unpleasing, but it is hardly systematic. What he didn't do was employ a clearly stated method to create a concise but effective core vocabulary, as Ogden (1937) and Hogben (1943) pointed out long ago. Most subsequent projects are open to the same criticism.

When it comes to creating a vocabulary, constructed languages take one of two main approaches:

- **Eclectic**, where the designers pick from a variety of linguistic sources, sometimes with a small admixture of completely made-up items. Examples include: Esperanto, Novial, Loglan, Unish.
- **Coherent**, where the vocabulary is drawn predominantly from a single source. Examples include: Latino sine Flexione (from Latin), Interglossa (from Greek), Interlingua (from the Romance languages), Lingua Franca Nova (from the Romance languages, apparently using Catalan as a kind of tie-breaker).

With the notable exception of Hogben's Interglossa (1943), none of these projects paid much attention to word economy, i.e. to establishing a minimal necessary core vocabulary. Indeed, the Interlingua English Dictionary (IALA, 1951) boasts of having 27,000 entries; while the Unish website (www.unish.org) has a section soliciting suggested new words from interested readers. In other cases the designers appear to have relied on their intuitions to decide how many and which words were necessary.

A Manifesto for Vegetarians, Esperantists & Other Cute Animals

My contention is twofold: firstly, that the world does need an international language; secondly, that it is possible to create a language that is superior for this purpose, in terms of learnability and usability, than either English or Esperanto.

1. **Orthography**: it is very easy to improve on English in this aspect, and not difficult to improve on Esperanto, where the accented consonants are an irritant. Several projects have already shown this, e.g. Lingua Franca Nova.
2. **Grammar**: English grammar is a minefield for the unwary, and Esperanto also contains some unnecessary pitfalls. Again, ways of improving on this have already been demonstrated by Lingua Franca Nova among other projects.
3. **Lexis**: Esperanto vocabulary is too large and disorderly, English much more so.

It is the third item that is really crucial, and that is where all previous projects have fallen down. I believe the time is ripe for a more systematic approach, with the aid of computer processing.

In this area we can take advantage of previous efforts to codify what Hogben calls a "list of essential semantic elements" (LESU). I know of four serious attempts at this task, from which a consensus can be obtained automatically (Ogden, 1937; Hogben, 1963; Macmillan, 2002; Longman, 2003).

The important novelty comes with simple a procedure for 'averaging' words, which I have implemented in the Python3 programming language.

Once we have our LESU -- in English to begin with (pardon my Anglocentrism) though the process could, and should, be repeated with other languages -- we can obtain translations in a number of other languages, let's say Portuguese, Spanish, French, Italian and Romanian for the sake of argument. These are candidates for the interlanguage vocabulary.

The next step is to use a textual similarity score to pick, for each item in the LESU, which of these natural-language terms is closest to the others (thus an analogue of the median). For example, applying this procedure to the word for 'young' in the five languages mentioned above, plus Latin,

['jeune', 'joven', 'giovane', 'jovem', 'tinar', 'iuventus']

we find that

'joven'

is the most 'typical', according to this scoring system. So in this case the Spanish word is the 'verbal average'.

Then a small-scale search is performed on each 'average' term, by systematically deleting one character at a time and also by exchanging adjacent characters, to see if this increases the similarity score. If it does, the modified version is kept, otherwise the original item, and this becomes our international word. To continue the example from above, this procedure alters 'joven' to

'jove'

which gets a slightly higher similarity score with respect to the 6 exemplars than 'joven'. This illustrates that the result doesn't have to be present already in one of the source languages. (It is present in Catalan.)

Thus we have a computer-assisted methodology for creating an international core vocabulary.

1. Identify a list of essential semantic units. (Considerable spadework has been done in this area by lexicographers in recent years.)
2. Find translation equivalents in a number of existing languages for these items.
3. For each unit find the most 'typical' from among the alternatives according to an objective scoring system.
4. Optimize the 'typicality' by systematic character-based manipulations.

For the first time in interlinguistics the choice of vocabulary items can be based on repeatable objective criteria rather than the subjective judgement of a designer or group of designers. I have implemented one scoring system, which I believe has merit, to start this ball rolling. Needless to say, I do not insist that it is the best that could be done, which is why the program is made public here for potential improvements.

Prospectus

We now have the means for getting closer to Zamenhof's (and Hogben's) dream, a concise, workable and international vocabulary for international intercommunication. Ultimately, this will allow us to devote our energies to specifying the criteria that an international vocabulary should satisfy, leaving the computer to find words that fulfil those criteria.

I believe it is worth pursuing this line of research. If it does take us one step further along the path to a genuine global language, I hope readers will agree that Douglas Adams had his tongue firmly in his cheek (lingua in gena?) when he wrote:

"the poor Babel fish, by effectively removing all barriers to communication between different races and cultures, has caused more and bloodier wars than anything else in the history of creation." (Adams, 1979: 50)

And if he was wrong about the Babel fish, he could also be wrong in his assertion that white mice have been running the world. Surely it's those philoprogenitive fluffy bunnies that really rule the world?

References

- Adams, D. (1979). *The Hitch Hiker's Guide to the Galaxy*. London: Pan Books.
- Hogben, L. (1943). *Interglossa*. Harmondsworth: Penguin Books.
- Hogben, L. (1963). *Essential World English*. London: Michael Joseph Ltd.
- IALA (1951). *Interlingua-English Dictionary*. New York: Frederick Ungar Publishing Co.
- Lanchester, J. (2012). *Whoops! Why Everyone Owes Everyone and no one can Pay*. London: Penguin Books.
- Longman (2003). *Dictionary of Contemporary English*. Harlow: Pearson Educational Ltd.
- Macmillan (2002). *MacMillan English Dictionary for Advanced Learners*. Oxford: MacMillan Education.
- Ogden, C.K. (1937). *The ABC of Basic English*. London: Kegan, Paul, Trench, Trubner & Co. Ltd.

Appendix

Constructed Auxiliary Languages :

Year	Language	Surname	Forename(s)
1661	Universal Character	Dalgarno	George
1668	Real Character	Wilkins	Bishop
1699	Characteristica Universalis	Leibniz	Gottfried
1765	Nouvelle Langue	de Villeneuve	Faiguet
1866	Solresol	Sudre	Francois
1868	Universalglot	Pirro	Jean
1880	Volapuk	Schleyer	Martin
1886	Pasilingua	Steiner	Paul
1887	Bopal	de Max	Saint
1887	Esperanto	Zamenhof	Lazarus
1888	Lingua	Henderson	George
1888	Spelin	Bauer	Georg
1890	Mundolingue	Lott	Julius
1892	Latinesce	Henderson	George
1893	Balta	Dormoy	Emile
1893	Dil	Fieweger	Julius
1893	Orba	Guardiola	Jose
1896	Veltparl	von Arnim	Wilhelm
1899	Langue Bleu	Bollack	Leon
1902	Idiom Neutral	Rosenberger	Waldemar
1903	Latino sine Flexione	Peano	Giuseppe
1906	Ro	Foster	Edward
1907	Ido	de Beaufront	Louis
1913	Esperantido	de Saussure	Rene
1922	Occidental	de Wahl	Edgar
1928	Novial	Jespersen	Otto
1943	Interglossa	Hogben	Lancelot
1944	Mondial	Heimer	Helge
1951	Interlingua	Gode	Alexander
1957	Frater	Thai	Pham Xuan
1961	Loglan	Brown	James
1983	Uropi	Landais	Joel
1996	Unish	Jung	Young Hee
1998	Lingua Franca Nova	Boeree	George
2002	Mondlango	Yafu	He
2011	Angos	Wood	Benjamin