5SSEL026 – Language Construction
Lecture 7
Artificial Languages 1

The story of artificial languages is, to a large extent, the story of their creators. This week we will look at artificial languages created primarily to enhance communication. They were all created to fill an actual or perceived gap in the way natural languages work, and they can be measured by their success as communicative systems.

WHY MAKE ARTIFICIAL LANGUAGES (1)?
There are many reasons for creating artificial languages. Some are:
- **Because human language cannot encode some things efficiently.** Language is good at negotiating social, cultural and affective meaning; and it is less good at technical meaning. Modern technical skills, like mathematics and logic, require their own coding systems; and, while gestural communication, like music and deaf sign languages, are expressive in their selected channel (aural in the case of music, gestural-visual for sign languages), it is difficult to record their form in another channel (such as writing).
- **Because a neutral language is needed.** Esperanto and Volapük were both created to provide a simple international language. Both were invented in the 1880s, when it seemed that Britain was winning the colonial race, and English was winning the lingua franca race. In the past, Latin has provided a neutral language, at least in Europe; and nowadays ELF (English as a Lingua franca) probably has that role; but in the 1880s there was concern by many European powers that if English became the default lingua franca then they would be diplomatically disadvantaged.
- **Because human languages are seen as not logical.** For some reason, humans place a lot of value on logical thought and communication – despite the fact that human logic is a product of the way we think, and not the cause of it. Lojban has been developed to maintain a close mapping between our cognitive logic and our communication. Lojban is based on an earlier language, Loglan, which was designed to test the Sapir-Whorf hypothesis: would speaking logically force us to think logically?

MORE EFFECTIVE CODES
Language is just one of the ways we exchange information. As well as the oral-aural route, we regularly use the gestural-visual route and the gestural-tactile route. We also use smell and taste; but these are usually automatic or autonomic signals, and do not require attention or volition.

Some examples of codes which are not conventionally linguistic but which do carry meaning are:
- **Music:** this uses the oral-aural route (or gestural-aural if musical instruments are involved), but it often does not have a language element; if it does, it's called singing. Music can involve conventionalised negotiation toward shared meaning (e.g. antiphony, or call-and-response), but it is usually mono-directional – the sender and the receiver do not exchange roles.
- **Mathematics:** this is often preceded with the phrase, "the language of", but it is not conventionally linguistic. In particular, there is no negotiation toward meaning – many of the meanings in mathematics are predefined by international agreement and are non-negotiable. Mathematics is, however, an open-ended system which can build unlimited constructs from finite means. In this way it can express theories about the way the universe works as relationships between things or concepts in the universe; and we can express Einstein's famous formulation, that [[Energy] equals [Mass] multiplied by [[the speed of light] multiplied by [itself]]], as $e=mc^2$ in mathematical notation.
- **Graphs:** these are "mathematics gone to the pictures". They rely on the conventions of statistics (a subset of mathematics) to show data relationships visually. The can show relative sizes, rates of change, correlations and other relationships between items of data in a way which allows the reader to see correspondences that might otherwise be opaque. They are not conventionally linguistic, but they allow the reader to negotiate toward the sender's meaning.
- **Pictograms:** these are pictures which have conventional meaning as well as representative meaning: they both depict and signify. They are one of the earliest forms of writing, and seem to have been used at Göbekli Tepe, 11,000 years ago. They are still used today in ISO (International Standardisation Organisation) graphical sets, and as emoticons. A fully pictorial language is certainly possible, and to prove it Joe Hale has translated Alice in Wonderland into emoji.[1]
- **Logic:** like mathematics, this is often preceded with the phrase, "the language of", but it is not conventionally linguistic. Logic is interested in the truth or falsehood of utterances, and has an internal system for proving truth which precludes negotiation toward meaning. However, it weirdly shares a very illogical function with language: the initial "facts" on which a proof is built do not need to be any more rigorous than "agreed as true" (i.e. axiomatic); this allows logic to prove the existence of unicorns, for a given meaning of unicorn.
- **Other sounds:** there are some non-linguistic vocal sounds which have conventionalised meanings and can be part of a language dialogue. An example is the raspberry, sometimes represented as *prrbt*. You can decide whether sounds which are treated as non-linguistic in English are treated as linguistic in your language.

THE SEARCH FOR A NEW LINGUA FRANCA
In the 1600s the universality of Latin was disappearing, as Christian Europe continued its separation into Catholic (Latin-using) and non-Catholic (Latin-abhorring) factions. As a result, the search for a "perfect" international language to replace Latin generated a range of candidates. The most complete design for a "perfect language" was produced in 1668 by John Wilkins (1614–1672), but the design principles proved both too unwieldy and too idiosyncratic to produce a viable language. Wilkins' attempt to make a logically consistent language based on categorisation was just too alien to human cognition to be usable.

VOLAPÜK
In the late 1870s and early 1880s, Johann Schleyer designed the first artificial international language to be taken up by a community of users. He called this language Volapük, which means "for world-speaking" (Vol-a-pük; Transliteration: World-for-speak). The unfortunate coincidence of the English word for regurgitation being used to represent the word "speak" has meant that Volapük has never been popular in English-speaking areas.

By 1889 there were about a million Volapük users, mostly in Europe, and the third Volapük convention (in Paris) was conducted entirely in the language. However, by 1890 the community was beginning to fragment: Schleyer was insisting on proprietary rights over the language, and resisted all innovations introduced by the language's academy. He took the language away from the community, which fragmented into small groups supporting a range of successor languages, while most of the Volapük clubs switched allegiance to Esperanto.

After Schleyer died (in 1912), there was an attempt in the 1920s to relaunch a simplified version of the language as Volapük Nulik (New Volapük), but the world had moved on, and Volapük has now become a linguistic oddity rather than a useful communication system.

ESPERANTO
While Schleyer was working on Volapük in Baden in Germany, Ludwig Zamenhof was working on Esperanto in Russian Poland. This was the first artificial language with a following large enough to sustain it over an extended period of time. Zamenhof published the first description of the language in 1887, writing as Dr Esperanto (one who hopes), and this title was quickly adopted as the name for the language replacing Zamenhof's choice of *lingvo internacia*. Esperanto was

never devised as a perfect language, or even as a first language; in fact, it's role was initially seen as everyone's second language – it was to be an international communicative pidgin, but not really a full language. However, it has since grown from a code to full human language, as any code or pidgin that survives beyond one generation tends to do.

Esperanto has continued to be of interest to linguists, and has developed over its 130-year history. It now has about 2 million fluent speakers, with about 1,500 people for whom it is their first language. It is recognised in many organisations around the World, and has the dubious honour of being disliked by dictators. Nazi Germany condemned Esparantists to the death camps, making it a language of the Holocaust, and Joseph Stalin sent Esperantists to the gulags as foreign spies.

Esperanto is the most successful of the international languages by a large margin, and it has even begun to evolve irregular forms. This both shows that it has become a natural language, and that the original objective of creating a completely regular second-language code was never really going to happen.

IDO
The Ido project was officially founded in 1907 with the intention of reforming and simplifying Esperanto. For this reason the language is very similar to Esperanto, and its name is an Esperanto word meaning "offspring". Ido simplified the Esperanto alphabet, got rid of the accusative noun case, changed the plural suffix from -oj to -i, and attempted to make the lexis more logical. However, most Esperantists refused the changes, so the project set off on its own trajectory.

It had a brief period of popularity, but never came close to replacing – or even competing with – Esperanto. The death of its founder, Louis Couturat, in 1914 (at the very beginning of World War I) meant that all progress in developing the language was suspended until 1920. Otto Jespersen then championed Ido until 1928, when he published his own language, Novial, to replace it. Novial sank without trace.

There was (and still is) an Ido academy to approve new additions to the language, but they approve only a handful of new words each year – which cannot even keep up with technological development. Imagine how you would cope if your language had to wait a few years before an official term for "mobile phone" was approved. The current Ido community is about 200-strong; if it were a natural language we would describe it as critically endangered.

FOR NEUTRALITY
The purpose of all these languages was to create a neutral language which every person could use but no country could own. However, the ongoing conflicts between adherents of the different languages show that the role of languages in differentiating between groups is probably as strong as their role in facilitating communication. The intention behind these international languages was to create a language which has no army and navy, a neutral language which will enable all nations to work together in harmony, and a language in which no group had a "home language" advantage. Yet the trappings of statehood (flags, badges, ideologies, and intergroup competition) seem to be a part of how the community of speakers define themselves.

In addition, the international languages encounter the same problems as natural languages:

- **Languages change as populations, technology, society and cultures change.** This results either in negotiations toward new consensus forms (Esperanto) or in new forms being imposed (Volapük). The first can lead to dialectisation and daughter languages (such as Ido); the second can lead to disaffection and apostasy. Either way, the language community begins to break up.

- **Languages drift:** idiolects and idiogrammars can become fashionable, then acceptable, then standard.

The big problem for artificial international languages is that a de facto international language seems to have emerged from socio-economic factors: first the British Empire and later the American hegemony have ensured the dominance of English – to the point where international collaborations are often conducted in English, even when no first-language English-speakers are involved. (For instance, English was adopted in 2008 as the official language of ASEAN, despite being a non-native, minority language in the region.) It seems that English, a deeply mongrel and libertine language, may be sufficiently neutral to avoid "home team" advantage.

A MORE LOGICAL LANGUAGE
Loglan (Logical Language) was invented by James Cooke Brown to test the Sapir-Whorf hypothesis. This hypothesis proposes that differences in the organisation of different language systems can affect the organisation of an individual's cognition; or, to put it simply, that the system of language used affects the system of thought. To test this, Loglan was designed to restrict certain linguistic effects which are common in natural languages. It is very difficult to use vagueness or metaphor in Loglan, because meaning is strictly delimited and closely tied to form. Human language relies heavily on vagueness and metaphor to create new meanings, so this makes Loglan an unusual language, resistant to change. It has a simple structure which should make it easy to learn, but the unnatural semantic system makes it quite difficult to master.

For Brown, Loglan is a research tool, and he has attempted (unsuccessfully) to retain copyright over its language systems. He insisted on his proprietary rights over the language, and he resisted all attempts at innovation which were proposed by the Logical Language Group, the main group of Loglan users. It appears he learned from Schleyer's mistake over Volapük and, when called upon, he was able to repeat it almost exactly. As a consequence, the LLG redesigned the language from the bottom up, establishing the new vocabulary around the sounds used to represent particular concepts in a selection of natural languages. They called this new language, lojban (no initial capital letter).

Despite the new approach, lojban has the same basic problem as Loglan: it is too complex to be used as a natural language. However, it does have an active community which negotiates new words into meaning; and it has a "figurative expression" marker which allows some use of metaphor. It has a dedicated following – none of whom can speak it fluently, but some of whom can produce effective written translations of shorter texts. As exercises in proving or disproving the Sapir-Whorf hypothesis, both Loglan and lojban have been inconclusive; but as exercises in languages as cultural tools, both have been informative.

Other attempts to create logical languages have encountered similar problems. A recent contender was the language Ithkuil, created by John Quijada to be "maximally precise but also maximally concise". He recognises that it is a language which would never occur naturally, but he nevertheless spent over 30 years creating it as a hobby. The language has a small following, and it has undergone two further revisions since its first publication in 2004. A further revision seems to be currently underway. Ithkuil was brought to the attention of a wider audience by a New Yorker article in 2012 (https://www.newyorker.com/magazine/2012/12/24/utopian-for-beginners), but it remains a niche language.

LANGUAGES BY OTHER MEANS
Not all languages use the oral-aural mode of transmission, and there are many examples of human signalling codes created using the gestural-visual or gestural-somatic communication routes. In addition, chemical signalling is common in Nature, with insects in particular using olfactory communication systems. Humans use all of

these signalling modes as part of their extended communication system.

Some of these human signalling codes have been adapted as the primary mode of communication for some individuals, which has allowed them to become full languages. Examples of this transmutation are particularly common in deaf communities around the world, in which eight families of sign language – each with their own sub-languages and dialects – are now recognised across the globe.

Deaf sign languages are ordinary language but using other means. It is no more difficult – or easier – to translate from BSL or Ameslan to English (or vice versa) than it is to translate any other language into English. However, representing sign languages in script is very difficult. Several schemes have been tried, representing particular gestures with special symbols, but none have caught on; this is mainly because sign language is, much more than spoken language, multi-modal. Where facial movements and body posture provide supporting semantics for spoken language, in sign language they are formally meaningful, and sometimes even grammatical. Where speech-meaning can be deciphered from a stream of individual sounds, which can be converted into a stream of individual marks on paper, sign language requires a full visualisation of the signer. Fortunately, nowadays we have ubiquitous video, which has removed many of the problems for record-keeping of signed conversations.

---

[1] https://creators.vice.com/en_uk/article/ez5vd4/author-translates-all-of-alice-in-wonderland-into-emojis.

Deaf-blind languages are also ordinary languages by other means, but they are less formalised and rely more on tactile letter-spelling than sign languages. However, this does mean that communication is easily encoded into a more permanent readable form, one version of which is Braille. Because many deaf-blind people have only become doubly disadvantaged in old age, each person tends to have their own ways of dealing with communication.

Other human gestural-visual communication usually involves letter spelling by an agreed convention. Systems like morse and semaphore are really just letter-replacement codes which work in any language which uses the Roman alphabet. The international Naval flag signalling system, in contrast, has only ten flags representing the numbers 0 to 9. These are hoisted in sets to represent multi-digit numbers which can then be checked against a code book to find their meaning. The advantage is that the flags themselves do not need to be a secret; if the opposing fleet does not have a copy of the code book, all they will get from the signal is a series of numbers.

When creating your language, remember that the written form is, for you, paramount. It is certainly possible to encode sound or gesture (or light displays) into your language, but you need to consider how they will be represented in text. All of these have been used in constructed languages in the past (some designed for this module), and even smell can be encoded with a bit of imagination. If you wish to use non-oral-aural communication channels in your language, feel free to discuss it with the lecturer.