5SSEL026 – Language Creation
Lecture 2
Continuous Creation

Languages do not emerge fully formed, they have origins; and languages do not stand still, they have histories and futures. So there are two big questions in this module: how does a new language come into being; and how does it develop? This lecture is about how languages change, and how we define a language.

WHAT COUNTS AS A LANGUAGE?

Every person has their own language. Sometimes two people speak in such a way that they can understand each other, at which point we say they are speaking the same language. But what counts as "same" and "different"?

- Mutually incomprehensible dialects can be treated as officially the same language, although they are, for practical purposes, different. Examples of this are Schwabian (German), Occitan (French), and Liverpudlian (English). Until relatively late in the 1970's, Catalan was viewed officially as a dialect of Spanish; but with the overthrow of Franco it was reassessed as a language in its own right.
- Languages can be treated as the same even though they have mutually incomprehensible lexes. Lexical problems can occur even where the languages are very similar, such as American and British Englishes (e.g. the two different meanings of *table a motion*: bring it up for discussion and decision in England, end discussion and postpone the decision in the USA).
- There are also mutually comprehensible languages (cognate languages) which are nonetheless seen as different official languages. Examples are the Norse group of Norwegian, Danish and Swedish; and the Serbo-Croatian group of Serbian, Croatian, Bosnian, Montenegrin and Kosovan. Max Weinreich is supposed to have said that *a language is a dialect with an army and a navy*, a good description of cognate languages.
- Finally, there are blend languages, which incorporate features of more than one language. Examples are Alsatian, which includes features of French and German; and Scots English, which mostly uses English grammar but Scots vocabulary.

Several linguists have compared language to a game, usually chess. If you are interested in this metaphor, see the paper at: http://martinedwardes.me.uk/playing_language_games.pdf.

GRAMMATICALIZATION

Grammaticalization describes how languages change over time, both lexically and grammatically. Some examples of recent grammaticalization in English are the use of "was like" to replace "said", and the emphatic negative, which places "not" at the end of an utterance.

- Heine & Kuteva (2007) show how language could have developed from single word utterances to simple grammars, and then to more complex grammars, through grammaticalization. They see the noun-verb differentiation as the first step to grammatical language, and everything is built from there. Once there are nouns and verbs, a need for noun and verb qualifiers will grow; when two-argument language constructs appear there is a need for question forms, linkers, and negators; the capacity to focus on different arguments in a multi-argument construct requires variable word order or grammatical marking; tenses are needed to describe non-current events; and so on.
- Dixon (1997) takes a strategic approach to language change, seeing language change as cyclical. He describes grammaticalization in terms of typological change: **isolating** languages, in which each word is a single meaning-unit, tend to become **agglutinating** languages, in which words contain multiple but separable meanings; agglutinating languages tend to become **fusional** languages, in which single syllables can perform multiple meaning functions; and fusional languages tend to become isolating.
- Hopper & Traugott (1993) take a more tactical approach: language change is a product of pragmatic, semantic, morphological and phonological variance. It happens because individuals change their idiolect over a lifetime; and because there is misinterpretation between individuals, and especially between generations. Reanalysis, in which a form remains unchanged but the underlying interpretation changes, is one type of variance (e.g. the adjective *wicked* changed it semantic field, and then became used as an exclamative). Another type of variance is reuse, where a form changes but the underlying interpretation remains the same (e.g. the word *verbing*, a verb-class word created from a noun-class root).

Languages change for at least four reasons:

- **External influence.** For instance, Old English underwent a massive change due to the imposition of a new language (Norman French). Middle English is structurally simpler and lexically richer than both Old English and Norman French.
- **Internal subgrouping.** For instance, regional habits can become fixed into dialects, such as the pronunciation differences between the Gileadite shibboleth and the Ephraimite sibboleth (for the story, see https://www.biblegateway.com/passage/?search=Judges+12&version=NABRE).
- **Language drift.** For instance, from 1350 until about 1600 there was a progressive change in English long vowel sounds: they moved away from the continental Germanic sounds and became pronounced higher in the mouth. So, for instance, the pronunciation of "sea" changed from "say" to "see". This became known as the Great Vowel Shift.
- **Cultural constraints.** For instance, the need for a massive increase in lexical terms (to deal with new sciences and new technology) created an environment where words are reused as different word-types. Nouns are used as adjectives (*management consultant*), verbs as nouns (*the ask is too great*) and nouns as verbs (*I'm keying in the data*). These, however, are just modern instances of a very old habit in English usage.

What tends to stay the same in a language is limited to some very basic functions. For instance, verbs and nouns tend to be stable as separate roots upon which other meanings can be built (but some languages, like Finnish, can use noun affixes on verb roots); and most languages allow at least three-argument forms (although Pirahã does things differently).

TYPING LANGUAGES

Dixon's three language typologies represent one way of typing languages; there are others. Unfortunately, different typologies often use the same terms to mean different things. One alternative typology to that of Dixon divides languages into:

- **Isolating:** each word has only one morpheme. Both meaning and grammar are built with words.
- **Analytic:** there are no grammatical morphemes. Meaning, but not grammar, is built with affixes.
- **Synthetic:** there are grammatical and lexical morphemes. Both meaning and grammar are built with affixes.

Synthetic languages are then further divided into:

- **Agglutinative:** each affix has a single grammatical or lexical role and a single form.
- **Fusional:** affixes can have more than one grammatical or lexical role. The form of the affix can be variable.
- **Polysynthetic:** the language makes extensive use of affixation and can create single-word sentences, or holistic utterances.

- **Oligosynthetic:** like a polysynthetic language, but there are a limited number of affixes. There seem to be no natural oligosynthetic languages.

The problems with this approach (which is used in Wikipedia) are:
- As a definition of possible languages, it is incomplete. For instance, it does not differentiate between variable-affix and invariable-affix fusional languages.
- It doesn't type languages, it types constructs. Most languages use most or all of these typing strategies.

GENDER AND CASE
Many languages have attentional grammar features. These are functions which are not vital to comprehension but which seem to add something to the language users' experience. They may reflect a cultural approach to the world, or they may represent a way for a culture to differentiate itself from its neighbours. Two of these features are **Gender** and **Case**.

GENDER
Gender is often applied to nouns and pronouns, frequently to adjectives, sometimes to verbs and rarely to other word types. Different languages have different gender groupings; but many seem to be variations on a twelve-part division: masculine, neuter and feminine subdivided into animate (human or nonhuman) and inanimate (concrete or abstract) (see lecture 1 handout).

Gender can also mark rank or deference, as in Spanish *usted/tú*, French *vous/tu*, and Middle English *you/thou*.

English is largely ungendered, which makes it difficult for native English speakers to understand why genders are used. That very incomprehension, though, illustrates the importance that gender has for in-group recognition. Non-native speakers of a gendered language make mistakes that identify them as non-native, simply because the gendering rules are not second nature to them. Genders are not just grammatical constructs, they are markers of group membership, and encapsulate important facts about the group's culture.

CASE
Case is a way of marking the roles of noun phrases within an utterance. In English the roles are indicated by position in the utterance (***Joan*** *gave the* ***man*** *a* ***book***), or by the adpositions used (*Joan gave a book* ***to*** *the man*), or by affixation (*Joan****'s*** *book*) or by lexis (*Joan* ***got to her house*** / *Joan* ***got home***). As with language typing, it is probably better to see these as strategies available within a language rather than as definitions of the language.

Case marking can occur just on nouns, or there can be agreement-marking on adjectives and even determiners. Eight noun cases are common in languages (although they do not necessarily all occur together in a single language):
- **Vocative:** a direct reference to the addressee of the utterance.
- **Nominative:** the subject of a finite verb (***Joan*** *took flowers)*.
- **Accusative:** the direct object of a transitive verb (*Joan took* ***flowers****)*.
- **Dative:** a general indirect object of a verb (*Joan stood* ***in the hospital***), or the receiver of the direct object of a transitive verb (*Joan took flowers* ***to the hospital***).
- **Ablative:** an indirect object which is distanced from the subject, or which delimits the subject (*Joan comes* ***from Kent***).
- **Genitive:** an indirect object which owns the subject or object (*the dreams* ***of Joan***; ***Joan's*** *dreams*).
- **Locative:** an indirect object which locates the action of the verb (Joan put the cat **in the box**).
- **Instrumental:** an indirect object which is used to perform the action of the verb (Joan closed the door **with alacrity**).

As with gender, the purpose of case seems to be more cultural than grammatical.

CONTACT LANGUAGES
Pidgins & creoles (contact languages) are social interlanguages in action. An interlanguage is a merging of two languages inside a single brain, and comes in two varieties: a **Pidgin** is an interlanguage negotiated between minds as the communicative activity progresses; a **Creole** is an interlanguage by history, and therefore has an established base of rules which can help to shortcut negotiation toward meaning. For this reason, creoles are often treated as full languages, while pidgins are usually treated as coding systems.

A Pidgin is a negotiated, temporary interlanguage, and has no first-language speakers; but can persist across generations if no permanent Pidgin-based community appears. It has:
- **Limited lexis**, usually borrowed from both (or many) source languages.
- **Simple grammar**; sometimes a pidgin will only allow two-argument utterances, thus limiting grammatical hierarchy in the pidgin.
- **Highly standardised forms**, often with no rhetorical or textual devices, such as the passive.

A Creole is a full language, and has a population who speak it as a first language. It is a second- or later-generation interlanguage, having features of both (or all) its source languages, but also having features specific to itself. It has:
- **As much lexis as is needed**, with the capacity to readily incorporate new meanings into existing frameworks.
- **Complex grammar**, allowing multiple-argument forms, hierarchy, recursion, regularity and irregularity.
- **As many rhetorical/textual devices as are required** by the speaker community.

Derek Bickerton and **Salikoko Mufwene** provide two alternative views of pidgins and creoles. They differ about the effects of several linguistic features on creoles and pidgins:
- Are creoles and pidgins different **states** or different **degrees** of language creation (i.e. are there fundamental differences between the ways the two types of language work)? Bickerton says they are different states, a creole is fundamentally different from a pidgin; Mufwene says they are different degrees, with creoles behaving more like full languages and pidgins less like them.
- Do creoles form a particular type of language with a distinct structure? Bickerton says yes, a creole evolves out of a pidgin, but it remains distinguishable from a full language; Mufwene says no, all forms of language are language, they exist on a cline from simple pidgins to esoteric full languages.
- Is a creole an inevitable and natural result of children being raised in a pidgin language community? Bickerton says yes, the linguistic needs of the child are partially met by the available communication system, which is then mutated by the children's Universal Grammar engine into a human language; Mufwene says this process is not inevitable, a pidgin can remain a pidgin if no permanent L1-speaking community develops around it.
- Is pidginization a degrammaticised non-linguistic process, or does it contain its own grammar, although in simplified form? Bickerton says it's non-linguistic, and has to be "massaged" through the minds of children to become linguistic; Mufwene says it is already linguistic, but complexity will only appear if it becomes necessary.

Bickerton is a Formalist and Nativist, Mufwene is a Functionalist and undecided about innateness. Their choice of -ism dictates their viewpoint – and yours will do the same.

NEGOTIATION TOWARD MEANING

Negotiation toward meaning is produced by two back-channel processes. First, the sender is constantly checking receiver comprehension and adjusting their utterance accordingly; second, the receiver is constantly checking the message received against what they believe to be the sender's intention, and questioning dissonances. Negotiation toward meaning is a continuous interaction between the parties in the communication, and it relies on a pre-existing genetic willingness to do that negotiation.

To explain this genetic willingness to negotiate toward meaning, Thom Scott-Phillips has proposed a capacity to signal signalhood: as well as giving the signal (the informative function) the sender has to be able to model the intention of the receiver in relation to the signal, and has to be able to identify how the receiver will detect the signal (the communicative function). So, to ensure reception, a sender must signal that they are signalling. He describes this as:

> The receiver must believe [1] that the sender intends [2] that the receiver believes [3] that the sender intends [4] that the receiver believes [5] that X is true.

This is certainly the case when the receiver is not actively seeking meaning; but if the receiver is as driven to seek meaning as the sender is to signal it, the complexity can be shared between two minds. The sender must be aware that the signal has to accommodate the receiver; but, if the receiver wants to receive the signal, and has a back-channel to express incomprehension, then that is all the sender has to be aware of. Meaning is not a pre-existing thing travelling from one mind to another, it is a negotiation between minds.