A primer on constructed languages
Nathan Sanders
University of Toronto

I have always wanted to be able to express music and love and the things that I have felt in their own proper language—not like this, not like this with the procession of particular English verbs, adjectives, adverbs, nouns and prepositions that rolls before you now towards this full stop and the coming paragraph of yet more words.

—Stephen Fry, Moab is My Washpot (1997:78)

This chapter provides a tour through the nature and history of constructed languages, or conlangs for short. Conlangs, such as Esperanto and Klingon, are languages or partial languages that have been intentionally created by the conscious effort of individual humans, unlike languages like Estonian or Korean, which arose naturally and effortlessly among entire populations from the general subconscious human capacity for language. I begin in Section 1 by highlighting the most ancient notions of consciously controlling and creating language. Then in Section 2, I continue with discussion of the first true conlangs from the Middle Ages up through the Renaissance and Enlightenment, when language creation was guided largely by religious and philosophical concerns, though more creative uses began to appear in this time as well. In Section 3, I explore more recent history, when conlangs were believed to have practical purposes as lingua francas to bridge linguistic divides between communities. This period also saw the further development of conlanging as an art form, most notably by J. R. R. Tolkien, setting the stage for the current era of conlanging, discussed in Section 4, where conlanging’s status as art and even as a profession has been solidified. In Section 5, I discuss the role of conlanging as a tool for language revitalization. I conclude in Section 6 with a summary of important terms and concepts useful for understanding this primer and the other chapters in this volume. Much of the content in this chapter owes a great debt to various scholarly works with extensive discussion of conlangs and the history of conlanging, especially Knowlson 1975, Eco 1993, Okrent 2010, Adams 2011, and Stria 2016, as well as Jörg Rheimer’s (2007–2010) useful online overview of conlang history.

1. Ancient notions of language construction. The languages we ordinarily encounter in our daily lives are the result of hundreds of thousands of years of natural, unintentional evolution from the prelinguistic vocalisms and gestures that our Paleolithic ancestors likely used to convey basic information (Danesi 2004). Their rudimentary communication slowly evolved into increasingly complex systems, ultimately becoming true languages capable of expressing the full range of everyday human experience. These so-called natural languages, or natlangs for short, are intimately tied with who we are, not just as individual speakers or members of linguistic communities, but more deeply as human beings. The capacity for language is perhaps the single most distinctive property of our species. It is thus no surprise that we have sought to control and manipulate language beyond its natural evolution, as it allows us to exert control over our very identity.

Poetry is one such way that we impose our will on language, intentionally shaping it with structural patterns not typical of everyday use. The earliest recorded poetry dates back more than 4000 years, to the Pyramid Texts (Old Egyptian hymns, prayers, and spells carved into the burial chambers of the Saqqara pyramids circa 2400 BC; Erman 2012:1), not long after the invention of

(1)  
\[ \text{nig u}_4\text{-bi-ta la-ba-gál-la} \]
\[ \text{ki-sikil-tur úr dam-ma-na-ka še}_{10}\text{ nu-ub-dúr-re} \]

‘Something which has never occurred since time immemorial:
Didn’t the young girl fart in her husband’s lap?’

This kind of linguistic manipulation, both sacred and playful, eventually expanded beyond the confines of natlangs. We learned to bypass the slow, effortless evolution of natlangs to invent new languages that we could control in every respect, intentionally and actively using our individual imaginations to create something with which we had a fundamentally different relationship than with natlangs. These deliberate linguistic creations have been called by a variety of names (see Section 5), but by the 1990s, the term conlang (a blend of constructed language, coined by Danish linguist Otto Jespersen when he debuted his own conlang Novial in 1928) had gained currency among the conlang community through online discussions and the founding of the popular Conlang email list in 1991 (Peterson 2015:11). For simplicity of discussion, I use conlang (to refer to one of these invented languages) and its derived forms conlangering (the practice of language construction) and conlanger (someone who engages in conlanging).

Given the human predilection for manipulating language, some form of conlanging has surely existed as long as language itself. The concept of conlanging is perhaps first explicitly mentioned in writing by Athenaeus of Naucratis in his Deipnosophistae (circa AD 230). Although the examples he gives are not full languages with detailed grammars, they do represent a rudimentary type of conlang known as a naming language: a set of neologisms that can be used as replacement vocabulary or to refer to things that otherwise have no suitable name. Athenaeus writes of people who had created their own naming languages, such as Dionysus of Sicily, who Athenaeus reports had invented words like μένανδρος (mēnandros) ‘wait-man’ to be used in place of παρθένος (parthenos) ‘virgin’ and μενεκράτης (menekrátēs) ‘strong-stay’ for στύλος (stūlos) ‘pillar’ (Deipnosophistae Book III.54).

A few centuries later, the first apparent discussion of a full conlang appears in Irish myths about the origin of Gaelic, as described in Longarad’s grammar of Old Irish, Auaricept na n-Éces (presumed to be from the 7th century). According to legend, King Fénius Farsaid of Scythia journeyed to Mesopotamia in the aftermath of the Biblical “confusion of tongues”, in which God fragmented human language (believed to have originally been a single homogenous language) to prevent them from completing the Tower of Babel and reaching Heaven (Genesis 11:1–9). In some versions of the story, Fénius was already present in Babel, perhaps having even been one of the tower’s builders (van Hamel 1915, Carey 1990). With the help of an army of scholars, Fénius studied the post-confusion languages, and he (or by some accounts, his grandson Goidel Glas; MacKillop 2005) combined the best parts of these languages into a new more perfect language, Goïdelc (Gaelic), the original Irish language (Williams 2016). This story seems to be the first written record of the very concept of a human being creating a full language, an ability
formerly attributed solely to deities. The Tower of Babel story itself relies on this divine power, and this basic theme dates back at least to the Sumerian epic “Enmerkar and the Lord of Aratta” from the 21st century BC, in which the god Enki similarly imposed linguistic diversity on humans (Kramer 1968). Crucially, what distinguishes the Goidec myth from these earlier stories is that the language in question was supposedly created by a mere mortal, an indication that humans had begun to believe that they could have access to what had previously been exclusively a godly ability.

2. Conlangs from the Middle Ages through the Enlightenment.

2.1. Linguistic mysticism. The belief in the connection between language creation and divinity shaped many early conlangs. If gods created languages, what would it mean for a human to dabble in this enterprise? A common thread in this view was that language could reveal deeper insights into human thought and the nature of the universe itself, so designing a perfect language could reveal truths obscured by the “corruptions” of natlangs (Wilkins 1668:6), perhaps even getting us closer to knowing the first language handed down to mortals by the gods, if not the gods’ own divine language itself. In this section, I provide an overview of these early conlangs, with discussion of a few notable examples; see Knowlson 1975 and Eco 1993 for more information on conlanging in this time period.

Half a millennium after the first record of the notion that a mortal could create a full language in the legendary origins of Goidec, German abbess and polymath Hildegard of Bingen (1098–1179) created Lingua Ignota ‘unknown language’, which is popularly credited as the first true conlang. However, Lingua Ignota is not much more than a partial relexification (or relex for short) of Latin. A relex is a type of naming language used in conjunction with an existing language’s grammar, replacing the words or morphemes but keeping the underlying structure. In this case, Lingua Ignota had over 1000 neologisms to replace Latin words, such as loifol ‘people’ replacing populus, but otherwise, it did not appear to have its own grammar separate from Latin, and it retained Latin words where no equivalent in Lingua Ignota existed. In addition to the language, Hildegard also created Litteræ Ignotæ ‘unknown letters’, a constructed writing system or neography, which she used to represent Lingua Ignota (Figure 1).

![Figure 1. Litteræ Ignotæ (adapted from folio 464v of RheinMain University’s digitization of the Riesencodex, http://hlbrm.digitale-sammlungen.hebis.de/handschriften-hlbrm/content/pageview/450562, accessed 18 August 2018)](image-url)

Despite the comparative lack of grammatical depth in Lingua Ignota from the perspective of modern conlangs, Hildegard is revered as a key figure in the history of conlanging as the first genuine conlanger. Fitting its name, the true purpose of Lingua Ignota is unknown, though it is likely that it was some sort of sacred language, perhaps used to create solidarity among
Hildegard and her fellow nuns. Hildegard’s extensive writings, including a record of Lingua Ignota, are collected in the Riesencodex (also known as the Wiesbaden Codex after its long-time home in Museum Wiesbaden), an enormous tome compiled shortly after her death. See Higley 2007 for fuller discussion of Hildegard of Bingen and Lingua Ignota.

Though Lingua Ignota almost certainly had some sort of sacred purpose, there is no indication that Hildegard believed she had uncovered a language actually spoken by divine beings. The first attempt at such a discovery seems to have been by mathematician John Dee and occultist Edward Kelley in the 1580s, who claimed that angels revealed their language to them (this language is popularly referred to as Enochian). Adding to the verisimilitude of their “discovery” was the neography they invented for Enochian (Figure 2). Its complex organization and ornate glyphs gave the language an otherworldly feel that suited its supposedly angelic speakers, helping this language capture the imagination of believers and skeptics alike for centuries. See Laycock 1978 for more information about the history and structure of Enochian.

![Figure 2](image)

Figure 2. Enochian neography (adapted from Laycock 1978:37)

The decline in importance of Latin by the 1600s left the scholarly world feeling the need for a suitable replacement universal language (Stria 2016:50–51); the local natlangs that had supplanted Latin were deemed too unruly for this purpose (see Wilkins 1688:1ff for a typical scathing view of the “defects and imperfections” in natlangs). This search became the primary focus of conlanging in the 17th century, with the resulting conlangs often referred to as philosophical languages. These philosophical languages were intended to be perfect languages that would give precision and clarity to reasoned expression and thus, presumably, to thought itself. This belief in a deep causal relationship from language to thought remains a common folk belief to this day, though there is little evidence that there is any significant effect of language on thought (Gleitman and Papafragou 2005).

The earliest of these philosophical languages is likely Jean Douet’s “escriture universelle” (1627), and many more followed, including Francis Lodwick’s “common writing” (1647, 1652), Sir Thomas Urquhart’s Logopandectesion (1652, 1653), Cave Beck’s Universal Character (1657), and Philippe Labbé’s Lingua Universalis (1663). The most important of these philosophical languages is arguably George Dalgarno’s Ars Signorum (1661). Its system of categorizing semantic concepts was developed further into a complex hierarchy by John Wilkins (1668), whose work in turn helped inspire Peter Mark Roget’s now ubiquitous thesaurus (Roget 1853:xxiv–xxv). This style of conlanging continued through the 17th century but were all but abandoned in the next century, with only scattered examples in the 1700s (Knowlson 1975).

2.2. Early artistic languages. This era also saw the beginnings of the use of artistic languages, or artlangs for short. Unlike the mystical and philosophical languages that seek to uncover deeper connections to the universe and the human mind, artlangs serve a very different purpose. They are designed to suit some creative purpose, usually as flavourful adornment in a work of fiction. These early artlangs were usually little more than rudimentary naming languages, so that characters, places, and important items or concepts in a novel could have a realistically foreign name. Although little linguistic creativity was required at the time, these early artlangs
represented the first step in the evolution of conlanging into the serious art form it has become today.

One of the earliest artlangs is the language used in the country of Utopia in Sir Thomas More’s 1516 *De optimo reipublicae statu, deque nova insula Utopia, libellus vere aureus, nec minus salutaris quam festivus clarissimi disertissimique viri Thome Mori include civitatis Londinensis civis & vicecomitis*, more popularly simplified as *Utopia*, itself an invented word meaning ‘nowhere’, coined by More from the Greek *οὐ (ou) ‘not’ and τόπος (tόpos) ‘place’*. The Utopian language was more than a naming language, though still only a relex of Latin. The language appears only as a few isolated words in the main text (such as *Buthrescas* for a type of religious fanatic and *Cynemernes* for the first day of a month) and a four-line poem in the addendum written by More’s friend Peter Giles (né Pieter Gillis).

Other notable examples of early artlangs were sparse naming languages invented by François Rabelais and Francis Godwin. Rabelais’ languages were included in his five-novel series (beginning around 1532) about two giants named Gargantua and Pantagruel, while Godwin included his conlang Lunar in 1638 novel *The Man in the Moone: or A Discourse of a Voyage thither by Domingo Gonsales the speedy Messenger* (which is further notable for being the first known example of science fiction published in English). Jonathan Swift went beyond mere naming languages for his 1726 novel *Travels into Several Remote Nations of the World. In Four Parts. By Lemuel Gulliver, First a Surgeon, and then a Captain of several Ships* (a.k.a. *Gulliver’s Travels*), which featured dialogue from multiple conlangs (Brobdingnagian, Laputan, and Houyhnhnm).

2.3. *Linguistic deception*. Other relexes abounded in the Middle Ages and the Renaissance, though typically with no uniquely identifiable creators or starting point (giving them an intermediate status between true conlang and specialized evolution of a natlang) and with no concern for holy or artistic purposes, as this was also the era of *cryptolects*, secret argots used by certain subcultures who sought to disguise discussion of their activities (thieves, homosexuals, and others who engaged in behaviour that conflicted with law and/or social norms). These argots were partial relexes, with various, often incriminating, words and phrases replaced with other words, foreign borrowings, and neologisms. For example, *black-box* ‘lawyer’ and *rum-dubber* ‘experienced lockpicker’ are expressions in Pedlar’s French, a thieves’ cant used in England (documented in B. E. 1698). Despite the desired secrecy, these cryptolects were often well-known enough to feature in public discourse, literature, and dictionaries, with many words transitioning from secretive use to popular slang; for example, modern English *fence* ‘recipient of stolen goods’ originated in Pedlar’s French. Prominent examples of these cryptolects include the jargon of the Coquillards in France (exemplified by Villon 1489), Rotwelsche in Germany (documented in Hüttlin 1509), and forbesco in Italy (documented in Brocardo 1545).

Conlangs have been used to deceive in other ways. One of the most notorious and elaborate linguistic hoaxes in history was perpetuated in England in the early 1700s, by a (possible) Frenchman operating under the pseudonym George Psalmanazar, who pretended to be a Formosan (an aboriginal native of Taiwan). He became a minor celebrity with his lectures and 1704 book *An Historical and Geographical Description of Formosa*, regaling the public and academics alike with detailed, horrifying tales of his supposed homeland and his path from savagery to Christianity (Keevak 2004). A crucial part of Psalmanazar’s charade was a conlang he used as his “native” language, which convinced potential skeptics, as they found it to be “so regular and grammatical, as well as different from all others they knew, both with respect to the
words and idiom, that they gave it as their opinion that it must be a real language” (Psalmanazar 1764:180–181). However, Psalmanazar’s “Formosan” conlang had no actual relationship to any real Formosan languages of the Austronesian family, such as Atayal or Bunun. People were convinced by Psalmanazar because England at the time had little knowledge of East Asia at all, let alone the Formosan people or their languages. In essence, Formosa could be whatever Psalmanazar said it was, and he was safe from challenge. Later in life, his conscience weighed heavily upon him, and Psalmanazar came to regret his “wretched youthful years” (1764:5). As an apology to the world, he laid out many details of his early life and his Formosan hoax in his memoirs (published posthumously in 1764), though he characteristically took the secret of his birthplace and real name to his grave.

2.4. Early conlanging beyond Europe. The record of non-European conlangs during this time is sparse. The most well-documented early conlang from outside of Europe is the 16th century religious conlang Bâleybelen (a.k.a. Balaibalan), created by Turkish dervish Muḥīṭ Gūleşnī (1528–1606) (Koç 2005, Emre 2017). Unlike like Lingua Ignota and the European cryptolects, Bâleybelen is not just a relex, since it has its own constructed morphology and syntax, though it is still heavily inspired by Turkish, Persian, and Arabic (Bausani 1970). There are also some notable non-European conlangs invented or first recorded in the 20th century, such as Damin in Australia (Hale 1992), Eskayan in the Philippines (P. Kelly 2012), and Medefaidrin in Nigeria (Gibbon et al. 2010), which have many structural and functional parallels with Bâleybelen and the European conlangs of this era, so while they may not belong to together in time, they surely belong in spirit.

3. Early modern conlangs.

3.1. Connecting the world through language. A main focus of conlanging efforts after the decline of interest in philosophical languages was finding an ideal auxiliary language, or auxlang for short, which could be used as a practical lingua franca for people with different linguistic backgrounds. Most auxlangs of this period were intended for international or even world-wide use; an auxlang with such a large scope is sometimes referred to as an international auxiliary language (IAL). The most famous and successful of these is Ludwik Lejzer Zamenhof’s Esperanto (1887), but many notable auxlangs came before it. One such example highlighting the inventiveness of conlanging is François Sudre’s Solresol (created and publicized beginning in 1827 but not published until 1866 by his widow after his death), in which the syllables can be mapped to musical notes, colours, numbers, abstract lines, gestures, etc. (Figure 3), allowing the language to be communicated in different modalities (speech, music, art, etc.; Rose 2013). For example, the name of Solresol itself can be communicated as the spoken or written word solresol; the musical tune G-D-G; a colour block made up of light blue, then orange, then light blue; the number string 525; pointing to the tip of the ring finger, then to the pinky knuckle, then back to the tip of the ring finger; etc. Any set of seven distinct atomic units of any type can be used, giving Solresol incredible flexibility in how it is expressed.
Other auxlangs of this period include Joseph Schipfer’s Communicationssprache (1839), a simplified version of French; Jean Pirro’s Universalglot (1868), one of the first fully developed auxlangs to integrate linguistic features from multiple languages; and Johann Martin Schleyer’s Volapük (1880), which was the most successful auxlang until Esperanto supplanted it. The particular successes of Volapük and Esperanto spurred the creation of many other auxlangs, especially by those who sought to improve upon previous auxlangs. The first off-shoot of Esperanto was Jacob Braakman’s Mundolinco (1888), but the most successful was Ido, the result of a battle among Esperanto enthusiasts over whether Esperanto should be, or could even be allowed to be, improved. Presumably created by Louis Couturat and Louis de Beaufront (né Chevreaux), Ido was submitted anonymously in 1907 to the Délégation pour l’Adoption d’une Langue Auxiliaire Internationale, an organization founded by Couturat and Léopold Leau and including in its membership the famed linguists Jan Baudouin de Courtenay and Otto Jespersen (who went on to publish his own conlang Novial in 1928). The Délégation’s support for Ido’s reforms, such as eliminating accent marks and case-marking, triggered a significant split in the Esperanto community (Forster 1982).

Attempts at auxlangs continued for much of the 20th century, though none managed to match Esperanto’s success, and Esperanto itself fell far short of Zamenhof’s intentions. A fascinating example is Charles Bliss’s sophisticated system of visual elements called Blissymbolics, developed beginning in 1942 and first published in 1949. Blissymbolics uses an intricate system of simplified iconic glyphs, notational conventions, and compositionality to allow for representation of a wide range of concepts; for example, ‘teacher’ and ‘student’ are each represented as a string of symbols, one for ‘person’, then one for either ‘give’ or ‘get’ as appropriate (they differ iconically by reversing the direction of the arrow), and finally one for ‘knowledge’, which is itself a blend of the symbols for ‘house’ and ‘mind’ (Figure 4). Blissymbolics failed to become the international auxlang Bliss had hoped it would be, but some organizations around the world have successfully adopted it as a communication method for people with speech disorders.

Figure 4. Sample words in Blissymbolics (adapted from the website for the Blissymbolics Communication Institute – Canada, http://blissymbolics.ca/, accessed 7 June 2018)
3.2. Rising artistry in conlangs. This period also saw authors taking more care in developing artlangs for their fictional universes. Like their earlier counterparts, artlangs in this period were largely naming languages, focused on vocabulary rather than syntax or morphology, a sharp counterpoint to the contemporaneous auxlangs. However, the quantity and quality of the constructed vocabulary of this time was noticeably improved over earlier artlangs, presaging an increasing level of rigour that conlangers would eventually embrace as standard practice. Notable examples include Frederick Spencer Oliver’s Poseid for his 1894 novel *A Dweller on Two Planets: Or, The Dividing of the Way*; Edgar Rice Burroughs’s Barsoomian for his John Carter of Mars stories, beginning with a 1912 magazine serial eventually re-published as the 1917 novel *A Princess of Mars* (Barsoomian was later expanded by Paul Frommer for the 2012 film *John Carter*); George Orwell’s Newspeak for his 1949 novel *1984*; Anthony Burgess’s Nadsat for his 1962 novel *A Clockwork Orange*; Frank Herbert’s Fremen for his Dune series of novels, beginning in 1965 with *Dune*; and Richard Adams’s Lapine for his 1972 novel *Watership Down*.

3.3. J. R. R. Tolkien. Among those who created artlangs in this time, John Ronald Reuel Tolkien was an aberration. The effort typically spent on a conlang by his contemporaries was proportional to its intended practical use: artlangs were minimalistic dashes of flavour added to a larger work of fiction, while auxlangs were intricate machines designed to support the full range of human communication. Tolkien bridged the gap between these two extremes by creating fully formed languages and giving serious, dedicated focus to every aspect of his languages, but without any larger functionality or purpose beyond the sheer intellectual joy of doing so. However, he believed that his “secret vice” would not be taken seriously on its own, so he wrote his Middle-earth novels as a way to showcase them. Thus, while other writers created conlangs for their fiction, Tolkien created fiction for his conlangs (Tolkien 1981:219). As a scholar of philology and practitioner of multiple languages, Tolkien drew heavily from his extensive linguistic knowledge in building his conlangs, setting a high standard in rigor and verisimilitude that would not be reached again until the closing decades of the 20th century. Tolkien did not just impact the art form of modern conlanging; he essentially defined it.

Tolkien showcased multiple conlangs in his Middle-earth novels, from *The Hobbit* in 1937 to his posthumous *The Silmarillion* in 1977: three languages for elves (Quenya, Sindarin, Old Entish, though only the single word *a-lalla-lalla-rumba-kamanda-lindor-burúme* ‘hill’ appears in Tolkien’s novels); three for dwarves (Khuzdûl), spirits (Valarín), and ents (one each for Rohirric and Westron); one for Volván (Old Entish, though only the single word *a-lalla-lalla-rumba-kamanda-lindor-burúme* ‘hill’ appears in Tolkien’s novels); plus the Black Speech, the language used for the One Ring’s inscription (2).

(2)  
Ash nazg durbatulûk, ash nazg gimbatul,  
ash nazg thrakatulûk agh burzum-ishi krimpatul.  
‘One Ring to rule them all, One Ring to find them,  
One Ring to bring them all and in the darkness bind them.’ (Tolkien 1954:247)

Interestingly, within the lore of the Middle-earth stories, the Black Speech was created by Sauron to unite his servants in Mordor, making it one of the most notable examples of a conlang designed to be understood within its associated fictional setting as an actual conlang, rather than as a natlang.
4. Increasing sophistication and community in modern conlanging. Tolkien did not live to see the true impact that his work would have on the evolution of conlanging. In 1974, a year after Tolkien’s death, the television series *Land of the Lost* premiered, and it featured the conlang Paku (sometimes called Pakuni), created by renowned UCLA linguistics professor Victoria Fromkin. Since she was paid to create the conlang for someone else’s purposes rather than her own, Fromkin is likely the first professional conlanger (*Bucks County Courier Times* 1974, Erickson 1998:114–115, Peterson 2015:11, Fimi and Higgins 2017:26). Though Paku is not as extensive as some of Tolkien’s more robust conlangs, Fromkin imbued it with her expertise in linguistics, giving it realistic structure and substance inspired by languages of West Africa, further leading the art of conlanging down the path started by Tolkien. Though inexpert conlangs would continue to be built, Tolkien, Fromkin, and those who followed in their footsteps were establishing a new normal, and in the ensuing decades, it became standard and expected for conlangs to be based on systematic principles and typological facts discovered by the blossoming science of linguistics.

A decade after Fromkin created Paku, linguist and science fiction author Suzette Haden Elgin created Láadan for her *Native Tongue* series of novels (beginning with *Native Tongue* in 1984). Like Fromkin, Elgin was a linguistics professor, and she similarly used her knowledge of linguistics to give Láadan naturalistic structure. Láadan also holds a deeper purpose as an *engineered language*, or *engelang* for short: a conlang designed to explore one or more specific linguistic properties. In this case, Elgin constructed Láadan to be a feminist language that directly encodes issues and perspectives relevant to women (Elgin 1988:3–6). In the same year, *Star Trek III: The Search for Spock* showcased the alien language Klingon, which went on to become one of the most celebrated conlangs of all time. A few words of Klingon were first created by actor James Doohan (who portrayed the character Lt. Commander Montgomery Scott, better known as Scotty) for 1979’s *Star Trek: The Motion Picture*. Marc Okrand (who, like Fromkin and Elgin, has a PhD in linguistics) was later tasked by Paramount to create a full language compatible with Doohan’s original words (Okrand et al. 2011, Marc Okrand p.c. December 2016). The massive success of Klingon marks an important turning point in the history of conlanging, bringing the path taken by Tolkien, Fromkin, and Elgin to a new level of prominence and respectability, with Okrand’s thoughtful, informed design setting a high bar for future conlangs.

This bar has been met many times since by conlangers, especially those with training in linguistics. Syntactician Matt Pearson created Hivespeak (*Thhtmaa*) for the 1996 television series *Dark Skies* while working on his PhD in linguistics (Matt Pearson p.c. September 2018); Paku creator Victoria Fromkin also created the vampire language for the 1998 film *Blade* (Conley and Cain 2006:21); Klingon creator Marc Okrand also created Atlantean for the 2001 film *Atlantis: The Lost Empire* (J. Kelly 2001, Marc Okrand p.c. December 2016); Said el-Gheithy, director of London’s Centre for African Language Learning, created Ku (Chi’itoboku) for the 2005 film *The Interpreter* (King 2005, Conley and Cain 2006:95); Paul Frommer (a communications professor with a PhD in linguistics) created Na’vi for the 2009 film *Avatar* and Barsoomian for the 2012 film *John Carter* (building upon Burroughs’s original work) (Ekman 2012, Fimi and Higgins 2017:26, Paul Frommer p.c. July 2017); linguistic anthropologist Christine Schreyer created Kryptonian for the 2013 film *Man of Steel*, Eltarian for the 2017 film *Power Rangers*, and Beama (Cro-Magnon) for the 2018 film *Alpha* (Christine Schreyer p.c. August 2018, Sanders and Schreyer 2019 (this volume)); linguist Nick Farmer created the Belter creole for the 2015 television series adaptation of James S. A. Corey’s *The Expanse* series of novels and related

In most of these cases, the conlangers in question have only worked on one or two conlangs as side projects from their regular careers. However, at least one conlanger has managed to shift this pursuit from a side project to full-fledged career. David J. Peterson (who studied linguistics as an undergraduate and for his masters) has long been an avid conlanger and active participant in the conlanging community, helping found the Language Creation Society in 2007. The creators of the 2011 *Game of Thrones* television series (based on George R. R. Martin’s *A Song of Ice and Fire* novel series, beginning with *A Game of Thrones* in 1996) wanted to have elaborated versions of the various languages used in Martin’s fictional world, so that extended dialogue in these languages could be spoken on screen. The creators turned to Arika Okrent, who directed them to the Language Creation Society. The Language Creation Society in turn collected and filtered a number of proposals from its membership and passed them along to the producers, who ultimately selected Peterson’s work. He rose to prominence for his versions of Dothraki and Valyrian and has built a prolific career as a conlanger, with a long list of projects, including work for the 2013 television series *Defiance*; the 2013 film *Thor: The Dark World*; the 2014 television series *Dominion*, *Penny Dreadful*, *Star-Crossed*, and *The 100*; the 2016 television series *The Shannara Chronicles*; the 2016 films *Doctor Strange* and *Warcraft*; the 2017 television series *Emerald City*; and the 2017 film *Bright* (David Peterson p.c. August 2018).

With so many high profile conlangs being built by conlangers trained in linguistics, conlanging is increasingly considered a serious activity requiring specialized knowledge, artistic skill, and hard work. It is now rarely sufficient for modern works of fiction to use impoverished conlangs in the antiquated model of haphazardly combining sounds and letters without an underlying structure. A notorious exception is the *Star Wars* franchise, which has generally avoided the kind of sophisticated conlangs used in the franchises listed above. Instead, the conlangs of the *Star Wars* universe are typically superficial impressionistic babbling used to approximate alien languages (Zimmer 2009), which can lead to odd results, such as in 2015’s *Star Wars: The Force Awakens*, when the character Tasu Leech says something sounding like *fi-dzi-ga-ni-ga-ma-di-ja* twice in the same conversation, once subtitled as ‘it’s over for you’ and the second time as ‘twice’. While the pronunciation is slightly different each time (especially the tone on the second syllable *dzi*, which is lower in the second occurrence), this would still be an unusual situation in a natlang (and thus, in a naturalistic conlang), though unsurprising given the origin of Leech’s speech, which is bits of Indonesian, Sundanese, and other Asian languages cobbled together by Sara Maria Forsberg (a Finnish singer who rose to fame on YouTube with her 2014 video “What Languages Sound Like to Foreigners”, in which she fluidly uses mostly spontaneous gibberish to mimic the sound of 20 natlangs; Zimmer 2016).

Though professional conlanging receives the most public attention due to the popularity of the franchises that feature conlangs, conlanging is increasingly being seen as a valid artistic pursuit with its own merits and purposes, aided in part by publication of conlang grammars, such as Okrand 1985/1992 and Okrand 1997 (for Klingon) and Peterson 2014 (for Dothraki). One of the earliest notable conlangs of this newly empowered generation of personal artlangs is Brithenig, created in 1996 by Andrew Smith as a linguistic thought experiment, to see what kind
of Romance language would result if Latin had displaced Celtic in the British Isles and had undergone Celtic linguistic change (Okrent 2009:288, Stria 2016:86).

Aided by the expansion of the internet in the 1990s and especially the creation of the Conlang email list in 1991, modern conlangers have developed a robust community for exchanging ideas, critiques, and tools, allowing them develop increasingly sophisticated and experimental conlangs. Many of these are catalogued and archived on various websites, including Linguifex (http://www.linguifex.com/) and the Conlang Wikia (http://conlang.wikia.com/), and many are discussed in the 2017 documentary film Conlanging: The Art of Crafting Tongues (http://conlangingfilm.com), directed by Britton Watkins. The Language Creation Society publishes an online journal, Fiat Lingua (http://fiatlingua.org/), and hosts the biennial Language Creation Conference (LCC). There are also countless conlang resources on the internet and in print, including important works like Rosenfelder 2010 and Peterson 2015. Long viewed as a fruitless pastime by linguists, conlanging has more recently even been the subject of serious academic study within linguistics itself, especially as a pedagogical tool (for example, Gobbo 2013, Weltman 2015, and Sanders 2016; various articles in Fiat Lingua, such as Schreyer 2013, Garrett 2016, Anderson et al. 2017, and Pearson 2017; a dedicated panel discussion entitled “Where do Conlangs Belong in the Academy” at LCC 7 held in 2017; and the Teaching Linguistics with Invented Languages organized session at the 91st annual meeting of the Linguistic Society of America in 2017, from which the contributions to this volume are largely derived).

5. Applications for language revitalization. Conlanging also has a practical role to play in efforts to revitalize endangered and even extinct languages. Once a generation of children do not acquire a language, its ordinary evolution via native acquisition is broken, which is a devastating blow that languages normally cannot recover from. As the remaining generations of native speakers age and die, the unconscious knowledge of the linguistic structure is lost. Languages can be documented, of course, but not matter how robust a corpus, grammar, or dictionary may be, even collectively, they are not a complete substitute for access to natively-acquired grammatical competence. Even for English, perhaps the most documented and rigorously analyzed language on earth, linguists are still discovering new facts and new data, so it is not hard to imagine how much information must be lost for typical cases of language documentation, which have vastly less material written in or about the language, often filtered through only a handful of speakers and linguists.

Although the natural development of a language may be irrevocably broken by the lack of intergenerational transmission, a sufficiently documented language may yet still be revitalized in some form, so that a version of the language may continue to be spoken and perhaps someday even be natively acquired again. Language revitalization is difficult, requiring not only extensive documentation and massive resources and time, but most importantly, motivation within the community. As Schreyer (2011) notes, language revitalization can take cues from conlanging to aid its success, by drawing upon some of the same technology and by tapping into the same “coolness factor” to build prestige and a stronger sense of identity.

Indeed, even the actual act of conlanging itself can be a part of language revitalization. A particularly fascinating case (reminiscent of the well-known creation of Modern Hebrew as a partly constructed revitalization of Biblical Hebrew; Spolsky and Cooper 1991) is that of Patxohã, a constructed form of Pataxó, an extinct Maxakalían language of the Macro-Jê family, formerly spoken by the Pataxó people in the southern part of the Brazilian state of Bahia. Anari
Braz Bonfim, a member of the Pataxó community and scholar of Pataxó culture and language, provides extensive information on the Pataxó people and their construction of Patxohã (Bomfim 2012, 2017), which I briefly summarize here. As part of a larger interest in preserving their cultural heritage, members of the Pataxó community began serious work on revitalizing Pataxó in the 1990s. They collected whatever records they could find, including the oldest known record of Pataxó, a list of 90 lexical items compiled by German Prince Maximilian of Wied-Neuwied in his famed 1815–1817 expedition to Brazil, documented in two volumes published in 1820 and 1821. By 1999, interest in revitalizing Pataxó reached a critical point, with increased interest from younger members of the community and the formation of the Projeto de Pesquisa e Documentação da Cultura e Língua Pataxó [Project for Research and Documentation of Pataxó Culture and Language].

The Patxohã revitalization effort is ongoing, and pieces of Pataxó have been collected from written records, songs, elders’ memories, related languages, and even remnants lurking in the Portuguese spoken by the Pataxó today. Patxohã now has over 2,000 carefully cultivated words and is being taught in schools. Crucially, this revitalization effort would not have worked without the strong motivation of the community and their willingness to engage in a bit of conlanging, by creating native-like neologisms (as seen in the constructed name Patxohã itself, which is a clever blend of Pataxó, atxohã ‘language’, and xõhã ‘warrior’) and by repurposing conflicting or redundant lexical items, so that every traceable word in the language could still retain value rather than being discarded. For example, both miãga and txiâng are documented with the meaning ‘water’, but txiâng has been repurposed with the narrower meaning ‘rainwater’. There are essentially no records of Pataxó grammar, so that has to be constructed as well, and the Pataxó researchers are basing Patxohã’s constructed grammar largely on Portuguese, because that is the current native language of the community, with some inspiration from the related language Maxakalí. Although the end result of this process will not be the original Pataxó (a fact explicitly recognized by the researchers in giving Patxohã a distinct name), it will still be a truly Pataxó construction, built by the Pataxó, for the Pataxó.

From religion and philosophy, to art and entertainment, to education and language revitalization, the human drive to build languages has a range of purposes, both fanciful and practical, and the various ways these disparate purposes connect to each other is worth cultivating and exploring.

6. Terminology summary. The following is intended to serve as a convenient reference for the reader of this collection, a catalogue of definitions of some of the most crucial terms related to conlanging that are used in typical discussions about conlanging. A natlang (natural language) is an ordinary language that evolved naturally without conscious planning, such as English, Finnish, Mandarin, Quechua, and Arabic. This stands in contrast to a conlang (constructed language), which is deliberately designed by one or more conlangers, the real humans whose imagination and hard work a conlang is derived from, such as Hildegard of Bingen, J. R. R. Tolkien, Marc Okrand, David J. Peterson, and the Pataxó community building Patxohã. A conlang may sometimes also be referred to as an artificial language, imaginary language, invented language, model language, or planned language, though these expressions may carry different connotations or even denotations for some people (van Oostendorp 2019). The English-speaking conlang community has largely accepted conlang as the preferred generic term, though the other terms do appear in some sources, especially prior to the online coalescence of the community in the 1990s. Some languages, such as sign languages like American Sign Language
and revitalized languages like Modern Hebrew, blur the line between natlangs and conlangs, because they have communities of native speakers like a natlang, despite having some amount of intentional design. Even an originally pure conlang like Esperanto has come to be natively acquired or nearly so by hundreds of speakers (Versteegh 1993, Corsetti 1996, Bergen 2001), allowing it to transition into having properties of a true natlang.

Conlangs can be classified most basically by whether or not they are modelled in large part on some existing language. An a posteriori conlang is a conlang primarily modelled on one or more specific existing languages, which may be natlangs or conlangs (for example, Herbert’s Fremen is based on Arabic, and Ido is based on the conlang Esperanto), while an a priori conlang like Klingon is created from scratch with no direct connection to any particular existing languages. Note that Klingon does have features in common with existing languages, such as the object-verb-subject word order it shares with Hixkaryana (Derbyshire 1977), but crucially, Klingon is not directly derived from Hixkaryana itself. The line between a posteriori and a priori is usually more obvious in the lexicon. With morphological and syntactic structure, the distinction is not as strict, because most conlangs use some grammatical features found in existing languages, if not directly borrowed, at least through inspiration or necessity. For example, all six of the possible orderings of subject, verb, and object are found in the world’s natlangs (Dreyer 2013), so any conlang with some ordering of those three units is necessarily duplicating the order from some natlang, whether intended or not. There is no objective turning point at which this kind of linguistic borrowing causes an a priori conlang to be classified as a posteriori instead, so the distinction is more of a gradient scale than a categorical contrast (van Oostendorp 2019).

In addition to its basic linguistic foundation or inspiration, a conlang may also be classified by its creator’s motivation. There are three major categories for why someone might want to create a conlang. Nowadays, the most popular type of conlang is almost certainly an artlang (artistic language), which is created for some predominately creative purpose, such as use in a work of fiction (for example, Klingon, Elvish, and Kryptonian) or to stand on its own as a work of art, perhaps even just for its creator’s own personal satisfaction. The creator of an artlang may also design or draw inspiration from a corresponding conworld (constructed world) and/or conculture (constructed culture) for the beings the conlang is intended for.

An auxlang (auxiliary language) is a conlang created for practical communication among groups of people with different languages; an auxlang intended for international or even worldwide use is sometimes referred to as an international auxlang (IAL). Experienced conlangers familiar with the history of conlanging are aware of the seemingly insurmountable obstacles preventing an auxlang from fulfilling its purpose, even on a relatively local scale, so they tend to eschew creating pure auxlangs. Indeed, the typically contentious ideological discussions surrounding auxlangs are one of the primary sources of tension within the conlang community; for example, the original Conlang email list eventually split in 1996 into two lists, Conlang and Auxlang, to keep the discussions separate (Peterson 2015:12).

An engelang (engineered language) is a conlang created to have some specific structural properties, such as a restricted grammar and/or lexicon (as in Newspeak), using a modality other than speech or sign (e.g. musical notes, colours, or textures), or violating certain linguistic universals (as in Sylvia Sotomayor’s (1980–2019) Kèlen). A prominent type of engelang is a loglang (logical language), which is created with some sort of overarching logical criteria, usually with the primary goal of having no ambiguity. The most notable well-developed
examples of loglangs are James Cooke Brown’s Loglan (1960) and its derivative Lojban, which was developed between 1987 and 1997 by the Logical Language Group (Cowan 1997).

These three categories, artlang, auxlang, and engelang, are not mutually exclusive, so a conlang may fall into more than one, as with Elgin’s Láadan, which was featured in her Native Tongue series of novels, and thus, is an artlang, but which was also designed with particular structural constraints in mind (having a distinctly female worldview), making it an engelang as well.

A naming language is a special type of minimal utilitarian conlang, usually an artlang for works of fiction, created primarily to account for a relatively small number of words and phrases, especially proper names and common nouns (as with Lapine). Despite the lack of explicit grammar, naming languages may still have an underlying set of implicit phonotactic and/or morphological rules governing what makes a well-formed word, though they may be somewhat more haphazard and fluid than in a fully specified conlang.

A relex (relexification) is an a posteriori conlang in which the lexicon of the source language is replaced (in whole or in part), while its syntax is retained. If the replacement vocabulary also comes from an existing language (making the relex doubly a posteriori), the language providing the replacement vocabulary is called the lexifier. Modern relexes are often disdained in the conlang community for being perceived as unimaginative and lazy (though older relexes, especially those marking notable milestones in the history of conlanging, such as Lingua Ignota, are treated somewhat more generously).

A cryptolect is a special way of speaking used to disguise a language, often in the form of a relex, with a variety of obscure slang, borrowings, and neologisms replacing important words. A cryptolect is arguably not so much a conlang as it is a specific register, though there is a certain amount of conscious intentionality to a cryptolect that distinguishes it from how registers normally evolve. A particular type of cryptolect is a ludling, which involves regular phonological, rather than lexical, manipulation of an existing language, so that the original language is disguised but recoverable by someone knowledgeable in the rules of the ludling (Laycock 1972). Some phonological manipulations found in ludlings include: moving phonemes, as with the baliktád ludling in Tagalog, in which salá:mat ‘thanks’ becomes tamá:las by reversing all of the phonemes (Conklin 1956); inserting phonemes, as with Jerigonza in Peruvian Spanish, in which maestro becomes cha-ma-cha-es-cha-tro by insertion of cha- before every syllable (Piñeros 1988); and combining both movement and insertion, as with Utrovački in Serbian, in which grad ‘town’ becomes ud za granje by reversing the order of gra- and -d and then inserting u- at the beginning, za in the middle, and -nje at the end (Rizzolo 2006).

Though not a conlang in and of itself, a neography (literally ‘new-writing’) or conscript (constructed script) is a similar concept, an invented writing system. A special kind of neography is a pasigraphy, which is designed as a written representation of meaning only, without connection to any particular language or pronunciation (such as Blissymbols). Neographies are often created to represent a conlang (such as Tokein’s Tengwar scripts for Elvish), though neographies for natlangs also exist (such as Sequoyah’s syllabary for Cherokee and Sejong the Great’s hangul for Korean). The distinction between a neography and a truly natural writing system can be difficult to define, since natural writing systems typically have much more intentionality behind them than natlangs do, but known natural writings systems like the Greek alphabet and the Chinese logosyllabary evolved in a largely organic and gradual way, rather than being intentionally designed as a complete writing system in a short period of time.
Acknowledgements

Many thanks to Paul Frommer, Khadija Jagani, Piers Kelly, Suzi Lima, Marc Okrand, Arika Okrent, Matt Pearson, David Peterson, Christine Schreyer, and Ida Stria for conversations and feedback which have helped improve this work in its depth, breadth, clarity, and accuracy.

References

B. E. 1698. A new dictionary of the terms ancient and modern of the canting crew, in its several tribes, of gypsies, beggers, thieves, cheats, &c., with an addition of some proverbs, phrases, figurative speeches, &c. Useful for all sorts of people, (especially foreigners) to secure their money and preserve their lives; besides very diverting and entertaining, being wholly new. London: W. Hawes, P. Gilbourne, and W. Davis.
Beck, Cave. 1657. The universal character by which all the nations in the world may understand one anothers conceptions, reading out of one common writing their own mother tongues. An invention of general use, the practice whereof may be attained in two hours space, observing the grammatical directions. Which character is so contrived, that it may be spoken as well as written. London: Thomas Maxey for William Weekley.
Bliss, Charles Kasiel. 1949. International semantography: A non-alphabetical symbol writing readable in all languages. A practical tool for general international communication, especially in science, industry, commerce, traffic, etc. and for semantical education, based on the principles of ideographic writing and chemical symbolism. Sydney: Institute of Semantography.


Lodwick, Francis. 1647. *A common writing: Whereby two, although not understanding one the others language, yet by the helpe thereof, may communicate their minds one to another*. London.

Lodwick, Francis. 1652. *The ground-work, or foundation laid, (or so intended) for the framing of a new perfect language: And an universall or commonwriting. And presented to the consideration of the learned, by a well-willer to learning*. London.


Oliver, Frederick Spencer. 1894. *A dweller on two planets: Or, the dividing of the way*. Los Angeles: Baumgardt.


Swift, Jonathan. 1726. *Travels into several remote nations of the world. In four parts. By Lemuel Gulliver, first a surgeon, and then a captain of several ships*. London: Benjamin Motte.


Urquhart, Thomas. 1653. *Logopandecteision, or, an introduction to the universal language digested into these six several books, Neaudethaumata, Chrestasbeia, Cleronomaporia, Chrysomystes, Nelcadicastes, & Philoponauxesis*. London: Giles Calvert and Richard Tomlius.


