ᚠᚢ·ᛒᚩᚦ·ᚠᚱᛟᚠᚱᚢᛟᚠ·ᚠᛚᛖᚱᛖᛚᛖ:ᚹᛁᚻᛟᚠᛟᛖᚻ·ᚦᚱᚠᚦᚢ·ᚠᚾᛞ·ᚹᛁᛏᛖᚾᚠ·ᚠᚱᚠᚠᚢᚱ·ᚠᚾᛞ·ᛇᚱᛚᚠ·ᚼᛖᚻᚦᚠᚾ·ᛏᛞᚾᚢᚻ·ᚠᚾᛞ·ᛏᚠᚻᛁᚾᛏ

A Grammar of the Dunsish Language

Ǫnë Bæŋŋsuúdaa á Huebni Fýyrþi

Kevin Wesley Sīnā

hæcceitas

A Grammar of the Dunsish Language

A Grammar of the Dunsish Language, 0th Edition

Last edited: January 28, 2018.

Front Cover Image:

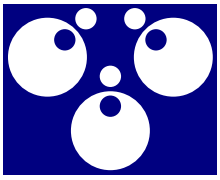
Unknown
Leaf from the Morgan Picture Bible (Scenes from the Life of Absalom), about 1250, Tempera colors, gold leaf, and ink on parchment
Leaf: 32.5 × 29.1 cm (12 $^{13}/_{16}$ × 11 $^{7}/_{16}$ in)
The J. Paul Getty Museum, Los Angeles

Back Cover Image:

Unknown
Adam Naming the Animals, about 1250–1260, Pen-and-ink drawings tinted with body color and translucent washes on parchment
Leaf: 21 × 15.7 cm (8 $^{1}/_{4}$ × 6 $^{3}/_{16}$ in)
The J. Paul Getty Museum, Los Angeles

Both the front and back cover illustrations are provided courtesy of the J. Paul Getty Museum Open Content Program. See `http://www.getty.edu/about/whatwedo/opencontent.html` for more information.

Designed and written with GNU Emacs.



Set in Linux Libertine, Anonymous Pro, Junicode, and Fira Sans with LuaLaTeX.

Epigraph translations are the author’s own, unless otherwise stated.

Published in New York, New York, United States.

Contents

List of Figures

List of Tables

ᚩ Os *byþ ordfruma   ǣlċre sprǣċe,*
ƿīsdōmes ƿraþu   and ƿitena frōfur,
and eorla ġehƿām   ēadnys and tōhiht.

Mouth is each language’s origin,
wisdom’s support and wise men’s comfort,
and every earl’s happiness and hope

Anglo-Saxon Rune Poem

𐌿𐌽𐍄𐌴 𐍈𐌰𐍂𐌾𐌰𐍄𐍉𐌷 𐍅𐌰𐌿𐍂𐌳𐌴 𐌰𐍄
𐌼𐌰𐌽𐌽𐌰𐌼 𐌹𐌽𐌽𐌿𐌼𐌰𐌽 𐌼𐌰𐌷𐍄 𐌹𐍃𐍄
𐌰𐌽𐌸𐌰𐍂𐌻𐌴𐌹𐌺𐌴𐌹𐌽 𐌹𐌽𐌼𐌰𐌹𐌳𐌾𐌰𐌽, 𐌹𐌸
𐌸𐍉 𐍅𐌴𐌹𐌷𐍉𐌽𐌰 𐍅𐌰𐌿𐍂𐍃𐍄𐍅𐌰,
𐌿𐌽𐌰𐌽𐌳𐍃𐌰𐌺𐌰𐌽𐌰 𐍅𐌹𐍃𐌰𐌽𐌳𐍉𐌽𐌰,
𐌲𐌰𐍃𐍅𐌹𐌺𐌿𐌽𐌸𐌾𐌰𐌽𐌳𐍉𐌽𐌰 𐌸𐌹𐍃
𐍅𐌰𐌿𐍂𐌺𐌾𐌰𐌽𐌳𐌹𐌽𐍃 𐌳𐍉𐌼

Unte ƕarjatoh waurde at
mannam innuman maht
ist anþarleikein inmaidjan,
iþ þo weihona waurstwa,
unandsakana wisandona,
gaswikunþjandona þis
waurkjandins dom

For all words taken from man can be transformed into another likeness, but these holy deeds, being unargued, make known the Maker’s judgement.

Skeireins VI:b–c

Foreward

Thank you for downloading my book! I hope you have as much fun reading it as I did writing it. Dunsish is a labor of love for me.

Acknowledgements

Thanks to everyone I'm thankful for!

Part I

Prolegomena

Chapter 1

Introduction

Dunsish is difficult to categorize. Conlangers generally recognize three basic categories:

1. Auxlangs, typified by Esperanto, designed to serve as an international auxiliary language.

2. Engelangs, or "engineered languages", designed to test some linguistic hypothesis, or reflect some logical structure. Lojban being preeminent.

3. Artlangs, with Tolkien's Quenya as the ur-example, which are constructed according to the creator's aesthetic taste.

Dunsish does not neatly fit into this classification system. While it is by no means intended to be any sort of auxilliary language, some of its design is indebted to the (at least, attempted) regularity and rich derivational morphology of Esperanto. Likewise, Dunsish is not strictly an engelang: it has "illogical" features which, although common in natural languages, do not serve any logical purpose beyond my own subjective taste for them. Nevertheless, Fythir does have some engelang-esque influences which would certainly be alien to any natural language. While Dunsish is primarily an artistic (so to speak) project, the term artlang usually implies some sort of naturalism, which Dunsish deliberately eschews (inasmuch as naturalism gets in the way of aesthetics).

Perhaps the closest language I know of (in design methodology, not actual grammar) is Jim Henry's wonderfully quirky language *gjâ-zym-byn* [1]. Both languages are personal projects, created solely for the edification of their designers. And likewise, neither make any attempt to restrict themselves to any particular branch of the conlang taxonomy.

Unlike (what I perceive as being) most conlangers, I have not made any attempt to attach any fictional culture or world to Dunsish, as such things do not interest me[1]. Accordingly, this book is written in a decidedly different manner than many, if not most, other conlang grammars. Since there is no "fourth-wall" (as it were) to break,

[1]Please note that I do not intend to disparage the work of conworlders here.

or internal history to describe, I have freely cited my influences for each aspect of Dunsish's design[2]. My hope is that these citations will help other conlangers who happen upon to this book in their quest to create a language of their own. Moreover, several features have been directly inspired by the work of my fellow conlangers, and I would not think it right to deprive them of credit for their insight.

[2]Barring, of course, those of my own invention.

References

[1] Jim Henry. *gjâ-zym-byn*. Dec. 2015. URL: http://jimhenry.conlang.org/gzb/gzb.htm (cit. on p. 3).

Chapter 2

Influences

In this section, I will list off my influences in creating Dunsish. All language data is taken from the Ethnologue [33], unless otherwise stated.

2.1 Natural Languages

African

- Afro-Asiatic:
  - Chadic:
    - Mwaghavul
    - Ron
  - Cushitic:
    - Beja
    - Somali
  - Egyptian:
    - Middle Egyptian
    - Coptic
  - Hamer
  - Semitic:
    - Amharic
    - Arabic
    - Akkadian
    - Hebrew
    - Maltese
    - Mehri
- Niger-Congo:
  - Balanta
  - Banyum
  - Bantu:
    - Bemba
    - Chaga
    - Chewa
    - Kukuya
    - Isu
    - Luganda
    - Swahili
    - Zulu
  - Grebo
  - Gbiri-Niragu
  - Edo
  - Ewe
  - Fula
  - Igbo

    - Janji
    - Phuthi
    - Supyire
- Nilotic:
    - Dinka
    - Mabaan
- Other/Isolates:
    - Fur
    - Kx'a
    - !Xóõ
    - Korana

Central American

- Uto-Aztecan:
    - Classical Nahuatl
    - Comanche
    - Hopi
    - Luiseño
    - Nahuatl
    - Southern Paiute
- Oto-Manguean:
    - Azoyú
    - Huautla Mazatec
    - Mixtec
    - Tlapanec
- Yuman:
    - Maricopa
    - Yavapai
- Other/Isolates:
    - Mixe
    - Tzeltal

East Asian

- Sino-Tibetan:
    - Belhare
    - Bantawa
    - Cantonese
    - Chepang
    - Classical Chinese
    - Classical Tibetan
    - Lhasa-Tibetan
    - Lhomi
    - Lepcha
    - Newar
    - Rai
    - Shanghainese
    - Shixing
    - Tibetan
- Japonic:
    - Japanese
    - Old Japanese
    - Ryukyuan
- Other/Isolates:
    - Kusunda
    - Mongolian
    - Thai
    - Vietnamese

Eurasian

- Caucasian:
  - Abaza
  - Abkhaz
  - Adyghe
  - Chechen
  - Georgian
  - Kabardian
  - Khinalug
  - Tsakhur
  - Tsez
  - Ubykh
- Dravidian:
  - Badaga
  - Malayalam
  - Tamil
  - Telugu
- Indo-European:
  - Ancient Greek
  - Bengali
  - Breton
  - Colognian
  - Elfdalian
  - Faroese
  - Farsi
  - Focurc
  - Galician
  - German
  - Gothic
  - Icelandic
  - Latin
  - Maldivian
  - Manx
  - Middle High German
  - Nepali
  - Old English
  - Old Irish
  - Old Norse
  - Polish
  - Punjabi
  - Romansch
  - Russian
  - Sanskrit
  - Scottish Gaelic
  - Serbian
  - Swedish
  - Tocharian
  - Welsh
  - Yola
- Paleosiberian:
  - Ainu
  - Aleut
  - Greenlandic
  - Inuktitut
  - Ket
  - Nivkh
  - Yukaghir
  - Yupik
- Turkic:
  - Karaim
  - Turkish
- Uralic:
  - Enets
  - Estonian
  - Finnish
  - Hungarian
  - Khanty
  - Livonian
  - Nenets

    - Saami
- Other/Isolates:
    - Basque
    - Burushaski
    - Elamite
    - Etruscan
    - Hattite
    - Hurrian
    - Jarawa
    - Sumerian

North American

- Algic:
    - Blackfoot
    - Ojibwe
    - Potawatami
    - Yurok
- Hokan:
    - Kashaya
    - Pomoan
    - Seri
    - Washo
    - Yana
    - Yuchi
- Iroquoian:
    - Cayuga
    - Proto-Iroquois
- Na-Dené:
    - Koyukon
    - Navajo
    - Tlingit
- Penutian:
    - Klamath
    - Mutsun
    - Nez Percé
    - Yowlumni
- Salish:
    - Halkomelem
    - Nuxálk
    - Saanich
- Wakashan:
    - Kwak'wala
    - Nootka
- Other/Isolates:
    - Haida
    - Hupa
    - Koasati
    - Kutenai
    - Lakota
    - Michif
    - Osage
    - Tanoan/Kiowa
    - Wichita

Oceanian

- Austronesian:
    - Acehnese
    - Gilbertese
    - Iaai
    - Kelabit
    - Lihir
    - Mela-Fila
    - Mwotlap

  - Rapa Nui
  - Rennellese
  - Rotuman
  - Sursurunga
- Pama–Nyungan:
  - Arrernte
  - Adnyamathanha
  - Dyirbal
  - Guugu Yimithirr
  - Kaytetye
  - Lardil
  - Walpiri
  - Wangkumara
- Trans-New Guinea:
  - Abui
  - Amele
  - Ekari
  - Kalam
  - Kobon
  - Kombai
  - Melpa
- Other/Isolates:
  - Jingulu
  - Kayardild
  - Manambu
  - Ngan'gityemerri
  - Semelai
  - Tok Pisin
  - Wagiman
  - Walman
  - Yele
  - Yimas

South American

- Arawakan:
  - Baniwa
  - Chamicuro
  - Tariana
- Barbacoan:
  - Awa Pit
  - Tsafiki
- Je–Tupi–Carib:
  - Guarani
  - Hixkaryana
  - Karajá
- Other/Isolates:
  - Aymara
  - Chiquitano
  - Movima
  - Quechua
  - Yaghan
  - Yagua
  - Warao

2.2 Constructed Languages

- Artlang:
  - Ayeri
  - Old Albic
  - Wessisc
  - Okuna
  - Moten
  - Himmaswa
  - Naduta
  - Idrani
  - Kahtsaai
  - Kgáweq'
  - Kēlen
  - Láadan
  - Lulani
  - Ŧhenqol
  - Elkarîl
  - Qþyn|gài
  - Quenya
  - Khafos
  - Valyrian
  - Miresua
  - Kirroŋa
  - Damin
  - Qevesa
  - Fukhian
- Loglang:
  - X-1
  - X-5
  - Plan B
  - Davin
  - gjâ-zym-byn
  - Liva
  - Lojban
  - Zango
  - Vorlin
  - Voksigid
  - Trari/Praet
  - Archeía/Verdeu
  - Sona
  - Konya
  - Ithkuil
  - Minyeva
  - Arahau
  - Kah
  - Ygyde
  - Kali-sise
  - Socialise
  - aUI
  - Blissymbols
  - AllNoun
  - Kalaba-X
  - Xaq
  - Bendeh
  - Ceqli
  - Gua\spi
  - FluidLang
  - Tilya
- Auxlang
  - Solresol
  - Esperanto
  - CosmicOS
  - Lincos
  - Poliespo
  - Ido
  - Volapük
  - Latejami
- Reconstruction:
  - Proto-Indo-European
  - Proto-Afro-Asiatic
  - Proto-Iroquoian

2.3 African

Mwaghavul

Language Facts	
Name	Mwaghavul
Other Name(s)	Mupun, Sera
Language Family	Afro-Asiatic, Chadic
No. Speakers	300,000
Area Spoken	Nigeria

Ron

Language Facts	
Name	Ron
Other Name(s)	Challa
Language Family	Afro-Asiatic, Chadic
No. Speakers	180,000
Area Spoken	Nigeria

Beja

Language Facts	
Name	Beja
Other Name(s)	Bedawi
Language Family	Afro-Asiatic, Cushitic
No. Speakers	1,200,000
Area Spoken	Sudan, Eritrea, Egypt

Somali

Language Facts	
Name	Somali
Native Name	*Af-Soomaali*
Language Family	Afro-Asiatic, Cushitic
No. Speakers	17,000,000
Area Spoken	Somalia

Middle Egyptian

Language Facts	
Name	Middle Egyptian
Other Name(s)	Classical Egyptian
Language Family	Afro-Asiatic, Egyptian
No. Speakers	*extinct*
Area Spoken	Egypt

Coptic

Language Facts	
Name	Coptic
Other Name(s)	Sahidic, Bohairic
Native Name	*Timetremənkhēmi*
Language Family	Afro-Asiatic, Egyptian
No. Speakers	*extinct*
Area Spoken	Egypt

Hamer

Language Facts	
Name	Hamer
Other Name(s)	Hamer-Banna
Language Family	Afro-Asiatic, Omotic
No. Speakers	74,000
Area Spoken	Ethiopia

Arabic

Language Facts	
Name	Arabic
Native Name	*al-ʿarabiyyah*
Language Family	Afro-Asiatic, Semitic
No. Speakers	290,000,000
Area Spoken	Arabia, Egypt, Levant

Akkadian

Language Facts

Name	Akkadian
Native Name	*akkadû*
Language Family	Afro-Asiatic, Semitic
No. Speakers	*extinct*
Area Spoken	Mesopotamia

Hebrew

Language Facts

Name	Hebrew
Native Name	*Ivrit*
Language Family	Afro-Asiatic, Semitic
No. Speakers	4,400,000
Area Spoken	Israel

Maltese

Language Facts

Name	Maltese
Native Name	*Malti*
Language Family	Afro-Asiatic, Semitic
No. Speakers	520,000
Area Spoken	Malta

Balanta

Language Facts

Name	Balanta
Other Name(s)	Balant, Balante
Language Family	Niger-Congo, Atlantic-Congo
No. Speakers	510,000
Area Spoken	Guinea-Bissau, Gambia, Senegal

Banyum

Language Facts

Name	Banyum
Other Name(s)	Banyun, Nyun, Bainouk
Language Family	Niger-Congo, Atlantic-Congo
No. Speakers	40,000
Area Spoken	Guinea-Bissau, Senegal

Bemba

Language Facts

Name	Bemba
Native Name	*Chibemba*
Language Family	Niger-Congo, Bantu
No. Speakers	4,100,000
Area Spoken	Zambia

Chaga

Language Facts

Name	Chaga
Other Name(s)	Kichagga
Language Family	Niger-Congo, Bantu
No. Speakers	1,300,000
Area Spoken	Tanzania

Chewa

Language Facts

Name	Chewa
Other Name(s)	Nyanja
Native Name	*Chichewa*
Language Family	Niger-Congo, Bantu
No. Speakers	12,000,000
Area Spoken	Zambia, Malawi, Mozambique, Zimbabwe

Luganda

Language Facts

Name	Luganda
Other Name(s)	Ganda
Native Name	*Oluganda*
Language Family	Niger-Congo, Bantu
No. Speakers	4,100,000
Area Spoken	Uganda

Swahili

Language Facts

Name	Swahili
Native Name	*Kiswahili*
Language Family	Niger-Congo, Bantu
No. Speakers	75,000,000
Area Spoken	East Africa

Grebo

Language Facts

Name	Grebo
Language Family	Niger-Congo, Kru
No. Speakers	390,000
Area Spoken	Liberia

Ewe

Language Facts

Name	Ewe
Native Name	*Èvegbe*
Language Family	Niger-Congo, Volta-Congo
No. Speakers	3,600,000
Area Spoken	Ghana, Togo

Fula

Language Facts

Name	Fula
Other Name(s)	Fulani
Native Name	*Fulfulde*
Language Family	Niger-Congo, Atlantic-Congo
No. Speakers	24,000,000
Area Spoken	West Africa

Igbo

Language Facts

Name	Igbo
Other Name(s)	Ibo
Native Name	*Asụsụ Igbo*
Language Family	Niger-Congo, Atlantic-Congo
No. Speakers	25,000,000
Area Spoken	Nigeria

Supyire

Language Facts

Name	Supyire
Other Name(s)	Suppire
Native Name	*Sùpyìré*
Language Family	Niger-Congo, Atlantic Congo
No. Speakers	460,000
Area Spoken	Mali, Ivory Coast

Dinka

Language Facts

Name	Dinka
Native Name	*Thuɔŋjäŋ*
Language Family	Nilo-Saharan, Nilotic
No. Speakers	1,400,000
Area Spoken	South Sudan

Apart from the Scandinavian languages, Dinka has had the greatest influence upon Dunsish's phonology and morphophonology. Dinka is notable for making a three-way phonemic vowel length contrast [24] [1].

Mabaan

Language Facts	
Name	Mabaan
Other Name(s)	Souther Burun
Language Family	Nilo-Saharan, Nilotic
No. Speakers	75,000
Area Spoken	Sudan

Fur

Language Facts	
Name	Fur
Other Name(s)	Konjara
Native Name	*Bèle Fòòr*
Language Family	Nilo-Saharan, Fur
No. Speakers	745,000
Area Spoken	Sudan, Chad

A notable feature of the Fur language is the extent to which metathesis forms a regular part of its morphophonology [11]. For example, when a consonantal personal prefix is attached to a verb stem, metathesis of the verb consonant and the following vowel is triggered. Hume-O'Haire et al. suggests that metathesis is best explained by a process of *perceptual optimization*, wherein the rearrangement of the verb root causes the personal prefix to be more phonologically emphasized and/or noticeable.

!Xóõ

Language Facts	
Name	!Xóõ
Other Name(s)	Taa
Native Name	*Taa ǂaan*
Language Family	Khoisan, Tuu
No. Speakers	2600
Area Spoken	Botswana, Namibia

2.4 Indo-European

Elfdalian

Language Facts

Name	Elfdalian
Other Name(s)	Övdalian
Native Name	*Övdalsk*
Language Family	Indo-European, North Germanic
No. Speakers	2000
Area Spoken	Älvdalen, Sweden

The primary interest of Elfdalian, with respect to Dunsish, is the considerable influence it has had on the phonology of the latter. In particular, Elfdalian preserves a number of archaic features of Old Norse lost in other Scandinavian languages. Moreover, it exhibits several innovative phonological processes which are typologically unusual in context.

Elfdalian notably preserves the Old Norse system of syllable length [31: p. 14]. Short syllables consist of a short vowel followed by a short, single consonant. Long syllables come in three varieties: a short vowel followed by a long consonant, a short vowel followed by a consonant cluster, or a long vowel followed by a single, short consonant. Finally, an overlong syllable consists of a long vowel followed by a long consonant[1].

The so-called "vowel balance" is a phonological process present in Elfdalian wherein the presence of a long root syllable causes the vowel of the final syllable of a word to to mutate [31: p. 16]. There are two primary examples: /i/ is transformed into /e/ and /å/ likewise becomes /a/. However, vowel balance is not applied universally. Some vowels and some endings are exempt from the process. In particular, enclitics are generally not subject to it.

Aside from the aforementioned system, Elfdalian also exhibits a rudimentary form of true vowel harmony [31: p. 17]. Vowel harmony appears to occur only in words consisting of short vowels. Elfdalian features a word tone system similar to Swedish [31: p. 14], and it is over this system which assimilation occurs. The harmony system causes any two consecutive short vowels to take so-called "level stress." As a result, both vowels have high tone and are equally stressed. Furthermore, in some cases, Old Norse words with multiple vowels have assimilated into the first of the two.

Elfdalian also exhibits complex sandhi and apocope [31: p. 17]. Sandhi phenonema are both word-external and word-internal (e.g. in compounds). Apocope involves the loss of the last vowel following a long or overlong syllable in a compound. Furthermore, words in a non-final position of a phrase are subject to the same process of apocopation.

[1] Sapir [31] appears to suggest that this list is not exhaustive.

Elfdalian is perhaps most famous for its complex vowel system which, save for a few exceptions, has a corresponding nasal vowel for each oral vowel [31: p. 21]. Nasal vowels form minimal pairs with oral vowels, thus nasality cannot be said to be merely allophonic. Nasality always occurs in vowels which precede a nasal consonant, but some nasal vowels occur before a since lost nasal consonant which was present in Proto-Germanic [31: p. 23]. The vowel system includes a large amount of diphthongs, some of which can occur is both long and short varieties. Moreover, it also includes two triphthongs: /juo/ and its nasal counterpart, both of which can be long or short. The presence of front nasal vowels is unique among European languages. As mentioned, most oral vowels have a unique nasal correspondent. However, there are two pairs of vowels which each share a single nasal vowel. In the case of /ɛ/ and /æ/, the nasal vowel is always realized as either /ɛ̃/ or /æ̃/, depending on dialect.

Faroese

Language Facts

Name	Faroese
Native Name	*Føroyskt*
Language Family	Indo-European, North Germanic
No. Speakers	66,000
Area Spoken	Faroe Islands

Along with the closesly related Icelandic, Faroese has contributed significantly to Dunsish's phonology and general aesthetic design goals. Notably, spoken Faroese distinguishes between alienable and inalienable possession [10].

Maldivian

Language Facts

Name	Maldivian
Other Name(s)	Dhivehi
Native Name	*Divehi-bas*
Language Family	Indo-European, Indo-Aryan
No. Speakers	340,000
Area Spoken	The Maldives

Maldivian, unique among Indo-European languages, and typologically unusual in its own right, features a duodecimal numeral system [6: pp. 107–126]. While the decimal system has largely replaced the older duodecimal system, there are still vestiges in the modern vernacular. For example, the Maldivian number *fasdoḷas*, sixty, literally translates as "five times twelve," or five dozens. Historically, the Maldivian islanders used a corresponding system of duodecimal-based weights and measures. It appears

that this number system is peculiar to Maldivian, as it is not attested in any related languages, e.g. Sinhalese. Table 2.1 lists Maldivian duodecimal number forms[2].

While there was almost certainly a period in which duodecimal was predominant in the Maldives, the ancestor of the decimal system used today is itself quite ancient [6: p. 108]. Thus it appears that both systems coexisted for quite some time, with the decimal historically assuming a marginal role. Strangely, it is the *decimal* numbers which show evidence of borrowing from Prakrits and/or Prakritisms.

The first ten numerals (in both systems) have differing attributive and substantive forms. The latter, which are shown in table 2.1, are formed by concatenation of the attributive root with the indefinite suffix *-e′* (with the underlying form *-ek*), which itself ultimately derives from the word for one. Only for numbers 11 and after do the duodecimal and decimal systems diverge. The duodecimal word for 11, *ekoḷos*, appears to ultimately derive from Proto-Indo-Aryan *ékādaśa*[3] [6: p. 122]. The decimal word for 11, *e(ň)gāra*, also ultimately derives from Proto-Indo-Aryan *ékādaśa*, albeit by way of Prakri *ek(k)ārasa* or *egārasa*, which of course explains the differences between the two words. Similarly, there is a duodecimal word for 12 (which exists in both Maldivian and Sinhalese) of a similar provenance: *doḷos*, which evidently comes from the Proto-Indo-Aryan *duvā́daśa*[4]. The origin of the decimal word for 12, *bāra*, is more difficult to establish. In either case, it appears that *bāra* is either a fairly recent Prakritism or borrowing [6: p. 111].

A major difference between the decimal and duodecimal systems are the regularity of numeral formation. By and large, the duodecimal numbers are the far more regular of the two. Decimal numbers in the range of 10—19 are irregular or have a seemingly irregular surface representation due to sandhi. One characteristic however, peculiar to both systems, requires further examination. Both duodecimal and decimal numbers which are one less than a multiple of their base (e.g. 35, 29) are formed in a unique manner: a prefix, *ekuni-* and *ona-* for duodecimal and decimal respectively, which means "one less" [6: pp. 122, 120].

Both diachronic and synchronic comparisons point to a considerable amount of reinterpretation of decimal numbers as numerically similar duodecimal numbers. The Maldivian word for 24, *fassihi* appears in an archaic form as *pasvisi* (cognate with Sinhalese *pasvisi*, 25). This is turn derives from Proto-Indo-Aryan *pañcaviṁśati* [6: p. 122]. The duodecimal number for 36, not being close to any multiple of 10, is simply *tin-doḷos*, or "three dozen." Duodecimal *fanas*, 48, likewise appears to be related to Sinhalese *panas*, 50, ultimately deriving from Proto-Indo-Aryan *pañcāśát* [6: p. 123]. Notably, the decimal and duodecimal systems, as mentioned, share a single word for 60, *fasdoḷas*, meaning "five dozen." In the decimal system, there is a secondary word for 60, *haṭṭi*, which derives from Proto-Indo-Aryan *ṣaṣṭí*, of the same meaning [6: p. 117]. Duodecimal *fāheti*, 72, is possibly a reinterpretation of Proto-Indo-Aryan *pañcasaptati*, 75. The duodecimal word for 84, *haddoḷos*, is similar to that of 36 and 60, meaning "seven dozen." Interestingly, *hia*, 96, is the largest number which may be expressed in the duodecimal system. As with the aforementioned examples, *hia*

[2]Morphemic representations (i.e. the underlying representation before being subjected to sandhi, etc.) are shown in between slashes. Alternative forms are shown inside angle brackets.

[3]literally, "one-ten" (cf. Latin *undecim*)

[4]literally, "two-ten" (cf. Latin *duodecim*)

was reinterpreted from its original meaning of 100 to the similar 96. Confusingly, in at least one instance a separate word for 100, *eksatēka*, was also used to refer to the value 96. Finally, it should be mentioned that *hia* appears also in a dialectical variation as the expanded *hia doḷos*.

Apart from two irregular words for "first," *furatama* and *paḷamu*, ordinals are regularly formed in Maldivian by means of the suffix *-vana* [6: p. 123]. Maldivian also has a series of collective numerals, which come in both inanimate and animate varieties [6: p. 124]. Inanimate collectives are formed by compounding with *eti*, "thing." For example, *dēti* /de-eti/, "pair, set of two" and *tin-eti*, "set of three." Conversely, animate collectives are formed in one of two ways, depending on dialect. Numerals may be compounded with *mīhun*[5], "people, men." In southern dialects, animate collectives are formed via the suffix *-verin*, which ultimate derives from the plural of a now obsolete noun, *veri*, which itself meant either "leader" or "person." Nevertheless, the two formations have slightly different semantics. Collectives formed with *mīhun* do not necessarily imply any sort of collectivity or cohesion betweens the members of the set, while those formed with *-verin* always do. Fractions are formed by the concatenation of the numerator, the word *bai*, meaning "part," the word *kuḷa*, preterite particle of the verb *kuranī* "to make, do," and lastly the denominator. Multiplicative numbers (e.g. twice, twofold, thrice, threefold) are formed by compounding with the word *guna*, "times" [6: p. 125]. There are three ways to express approximate values in Maldivian. Two numbers can be joined together, to express uncertainty across a range. Numeral substantives in the dative can also express approximate values [6: p. 126]. In some dialects, placing a collective form in the dative can be used to express greater uncertainty. However, dative approximates are limited to the numbers 60 and below. For numbers larger than 60, the postpositional adverb *varaka′*, "very," may be used.

Unfortunately, Fritz [6] is unclear as to whether duodecimal numbers can form ordinals, collectives, fractions, or multiplicatives. As fractions are expressed analytically, it is probable that duodecimal numbers could be used in such phrases. The use of the duodecimal *doḷas* in an example of an approximate number would seem to suggest that dative approximate constructions are possible with duodecimal numbers.

Tocharian

Language Facts

Name	Tocharian
Native Name	*Ārśi-käntu* [13]
Language Family	Indo-European
No. Speakers	*extinct*
Area Spoken	Xinjiang, China

[5] plural of *mīhā*, "person, man"

№	Duodecimal		Decimal	
1	*ekeʹ*	/ek-ek/	—	
2	*dēʹ*	/de-ek/	—	
3	*tineʹ*	/tin-ek/	—	
4	*hatareʹ*	/hatar-ek/	—	
5	*faheʹ*	/fase-ek/	—	
6	*haeʹ*	/ha-ek/	—	
7	*hateʹ*	/hat-ek/	—	
8	*aśeʹ*	/aś-ek/	—	
9	*nuvaeʹ*	/nuva-ek/	—	
10	*dihaeʹ*	/diha-ek/	—	
11	*ekoḷos*		*eṅgāra*	⟨*egāra*⟩
12	*doḷos*		*bāra*	
13	*doḷos ekeʹ*		*tēra*	
14	*doḷos dēʹ*		*sāda*	
15	*doḷos tineʹ*		*fanara*	
16	*doḷos hatareʹ*		*sōḷa*	
17	*doḷos faheʹ*		*satāra*	
18	*doḷos haeʹ*		*aśāra*	
19	*doḷos hateʹ*		*ona-vihi*	⟨*navāra*⟩
20	*doḷos aśeʹ*		*vihi*	
21	*doḷos nuvaeʹ*		*ekā-vīs*	
22	*doḷos dihaeʹ*		*bā-vīs*	
23	*ekuni-fassihi*		*tē-vīs*	
24	*fassihi*		*sau-vīs*	
25	*fassihi ekeʹ*		*fansa-vīs*	
26	*fassihi dēʹ*		*sabbīs*	
27	*fassihi tineʹ*		*hatā-vīs*	
28	*fassihi hatareʹ*		*aśāvīs*	
29	*fassihi faheʹ*		*ona-tirīs*	⟨*navā-vīs*⟩
30	*fassihi haeʹ*		*tirīs*	
31	*fassihi hateʹ*		*tirīs ekeʹ*	⟨*ettirīs*⟩
32	*fassihi aśeʹ*		*tirīs dēʹ*	⟨*battirīs*⟩
33	*fassihi nuvaeʹ*		*tirīs tineʹ*	⟨*tettirīs*⟩
34	*fassihi dihaeʹ*		*tirīs hatareʹ*	⟨*sauratirīs*⟩
35	*ekuni-tin-doḷos*		*tirīs faheʹ*	⟨*fansatirīs*⟩
36	*tin-doḷos*		*tirīs haeʹ*	⟨*sattirīs*⟩
40	*tin-doḷos hatareʹ*		*sāḷīs*	
48	*fanas*		*sāḷīs aśeʹ*	⟨*aśāḷīs*⟩
50	*fanas dēʹ*		*fansās*	
60	*fas-doḷos*		*fasdoḷas*	⟨*haṭṭi*⟩
70	*fas-doḷos dihaeʹ*		*haidiha*	⟨*haiteri*⟩
72	*fāheti*		*haidiha dēʹ*	⟨*bāhaiteri*⟩
80	*fāheti aśeʹ*		*aḍḍeha*	⟨*āhi*⟩
84	*haddoḷos*	/hat-doḷos/	*aḍḍeha haeʹ*	⟨*sayāhi*⟩
90	*haddoḷos haeʹ*		*nuva-diha*	⟨*navai*⟩
96	*hia*	⟨*hia doḷos*⟩	*nuva-diha haeʹ*	⟨*sayānavai*⟩
100	—		*satēka*	⟨*satta*⟩
200	—		*duisatta*	
300	—		*tin-satēka*	
1000	—		*hās*	
2000	—		*de hās*	

Table 2.1: Maldivian Duodecimal Numbers

By some accounts, Tocharian distinguished between a dual (used for casual pairs) and an ambal or "paral" (used for natural pairs, e.g. eyes, ears, etc.) [13]. However by some other accounts, this appears to be a mistake [12: p. 26]. Nevertheless, Tocharian featured a collective[6] suffix, *-aiwenta*, used to pluralize sets of objects [12: p. 26].

Tocharian had a highly innovative/divergent case system. Nouns inflect for 9 cases, but only three of these, nominative, accusative (or oblique), and genitive, ultimately stem from the Proto-Indo-European case system [12: p. 27]. On top of the fragment of its original, PIE-descended case system, Tocharian also developed a secondary agglutinative case system which diachronically resulted from the grammaticalization of postpositions [12: p. 28]. These six remaining secondary cases are the perlative, allative, comitative, ablative, locative, and instrumental[7] [12: p. 29]. It must be noted that these secondary case endings attach to the last word in a phrase, similar to the English possesive, and so might more accurately be termed clitics [12: p. 54]. Finally, Tocharian B features an unproductive "causal case," which is limited to use with only certain abstract nouns [12: p. 28].

Finally, the tocharian verbal system features an innovated system of forming causative stems through suffixation [12: p. 42].

2.5 Isolates

Burushaski

Language Facts

Name	Burushaski
Native Name	*Burū́šaskī*
Language Family	Isolate
No. Speakers	87,000
Area Spoken	Pakistan

Burushaski has an interesting gender system which recognizes four different classes: human males, human females, discrete non-human objects, and continuous non-human objects [2: p. 1233].

[6] termed "numerative" in some sources

[7] The instrumental case is found only in Tocharian A.

Chiquitano

Language Facts

Name	Chiquitano
Other Name(s)	Bésiro
Language Family	Isolate
No. Speakers	5900
Area Spoken	Bolivia

Chiquitano, also known as Bésiro, is a language isolate spoken by approximately 5,900 people in Bolivia [33]. Sans [30] presents evidence to suggest that Chiquitano has a quadripartite morphosyntactic allignment.

Movima

Language Facts

Name	Movima
Other Name(s)	Mobima
Native Name	*Chonsinelh*
Language Family	Isolate
No. Speakers	1452
Area Spoken	Bolivia

2.6 Papuan

Amele

Language Facts

Name	Amele
Other Name(s)	Anêm
Language Family	Trans-New Guinean, Madang
No. Speakers	5300
Area Spoken	Papua New Guinea

Amele, also known as Anêm, is a Madang language spoken by approximately 5,300 people [33]. Amele is notable for distinguishing 31 different possesive classes [14].

Ekari

Language Facts

Name	Ekari
Other Name(s)	Ekagi
Language Family	Trans-New Guinean
No. Speakers	100,000
Area Spoken	Papua New Guinea

Ekari is a Trans-New Guinea language spoken by approximately 100,000 people [33]. Uniquely, Ekari uses a sexagesimal (base-60) numbering system [4].

Gilbertese

Language Facts

Name	Gilbertese
Other Name(s)	Kiribati
Native Name	*Taetae ni Kiribati*
Language Family	Austronesian
No. Speakers	120,000
Area Spoken	Republic of Kiribati

Gilbertese phonology consists of a fairly standard 5 vowel system, and a (relatively) small inventory of 10 consonant phonemes. Length is contrastive for both vowels and nasal consonants [3: pp. 205–206]. Though Blevins et al. suggest that "evidence for the syllable as a phonological constituent in Gilbertese is not abundant," there are several facts which point to its validity. First, Gilbertese speakers evidently have no trouble in identifying syllable boundaries. Second, assimilation predictably fails to occur across such identified syllable boundaries [3: p. 207]. Gilbertese syllables are subject to seven constraints, show in figure 2.1.

Gilbertese is VOS [9: p. 325].

Tok Pisin

Language Facts

Name	Tok Pisin
Other Name(s)	New Guinea Pidgin
Language Family	Creole
No. Speakers	2,000,000
Area Spoken	Papua New Guinea

- ▶ 1 A syllable always has a nucleus.
- ▶ 2 The sonority of the nucleus is equal or higher to that of the onset.
- ▶ 3 Sonority continuously lowers after the nucleus.
- ▶ 4 Two adjacent morae with equal sonority only occur in the nucleus.
- ▶ 5 Onsets do not have clusters.
- ▶ 6 Rimes are composed of only vowels or nasal consonants.
- ▶ 7 Rime nasals which are neither word-final nor labial are homorganic with any following consonants.

Figure 2.1: Gilbertese Syllable Constraints

- ▶ 1 If sonority(x) ≥ 2, then x is a possible nucleus.
- ▶ 2 If sonority(x) ≤ 2, then x is a possible onset.

Figure 2.2: Gilbertese Sonority Rules

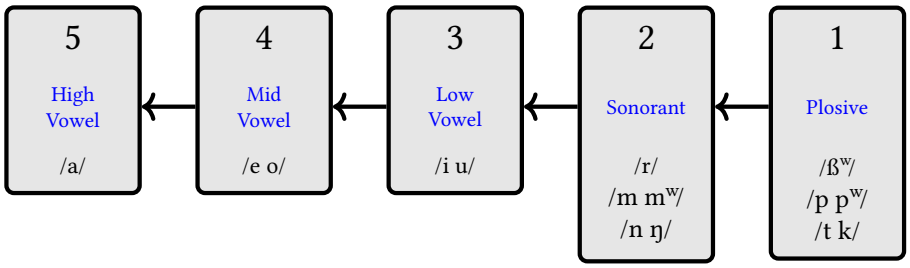


Figure 2.3: Gilbertese Sonority Scale

2.7 Central American

Seri

Mixe

Language Facts	
Name	Mixe
Native Name	*Ayuujk*
Language Family	Mixe–Zoquean
No. Speakers	90,000
Area Spoken	Mexico

The vowel system of Mixe is notoriously complex. Lowland Mixe has six different vowel qualities, on top of which three lengths and three kinds of glottalization are phonemically contrasted [7: p. 9].

Four types of basic stems can be distinguished: nominal, verb, adjectival, and adverbial. Furthermore, stems may be characterized according to their structure. Simple stems consist of a single basic root. Complex stems are those composed of a single root with one or more derivational affixes. Finally, compound stems have two or more basic roots, along with any derivational affixes [7: p. 20].

Verb stems come in a further two varieties: variable and invariable, which are distinguishable by syllable nucleus and coda. The former type undergoes various modifications of the nucleus (i.e. in the length or glottalization of the nucleus vowel) in certain environments [7: p. 21]. Verbs roots are either one or two syllables long. Two syllable verbal roots always have as their first syllable one of 22 different so-called "root formatives" of uncertain meaning [7: p. 23]. Unlike normal prefixes, however, formatives are retained even in compounds. Mixe has two main verbalizing suffixes, used to derive verbs from nominal roots. The first, *-ʔaHt*, creates stative verbs. The

second, *-i•y*, is of greater interest[8]. It is described by van Haitsma et al. as being "used to indicate something which requires some time to effect the condition, or a process" [7: p. 24].

Verb compounds are one of two types: specified-action or concomitant-action. Both have two "slots" which are restricted to certains sets of stems. For specified-action verbal compounds, the first and second slots are called the specifier and action slot, respectively. The former takes any nominal, adjectival, or adverbial root. The latter takes a verb stem, including stems which are themselves compound. Thus Mixe allows for recursion in compounding. The root in the specifier slot in some way modifies the stem in the action slot, however the precise relationship between the two is vague and semantically underspecified. Concomitant-action compounds have essentially two action slots, and implies some sort of concurrency or compositionality between both stems. While specified-action verbal compounds can be put in either of the slots of a concomitant-action compound, there is a limit of one; the other slot must contain a simple verbal root. Furthemore, concomitant-compounds cannot be nested. While many verbals stems can appear in either the first or second slot, others are restricted to one or the other [7: pp. 24–28].

Mixe features a wide array of derivational affixes that can be applied to verbal stems. Prefixes are either transitizing or intransitizing. Suffixes are split into two groups, and serve a broader range of functions. Table 2.2 gives an overview of some of the prefixes which are of particular interest [7: pp. 29–31].

	Type	Description
hu•-	Trans.	circumvent, go around (adverbial)
ko•-	Trans.	to act for a specific purpose
ku•y-	Trans.	to act in a specific manner or order
mɨ•(:d)-	Inverse	associative/conjunctive
ni•-	Trans.	purposive, to act for/about
nɨ''-	Trans.	pertaining to the body or skin (also used figuratively)
yaH-	Trans.	causative, polysemous, can promote locatives or instrumentals to D.O.
ʔa•-	Intrans.	iterative, pertaining to openings

Table 2.2: Mixe Verbal Prefixes

Table 2.3 lists suffixes which I have found noteworthy [7: pp. 33–36]. The presence of suffixes are among those conditions which can trigger the aforementioned stem nucleus mutations. While it is difficult to describe any precise semantic criterion which distinguishes either set of suffixes, those with meanings related to tense and person are found solely in the second group.

[8]In the orthography used by van Haitsma et al. [7], a capital 'H' denotes aspiration, the interpunct represents a mid-length vowel, a single tick (′) represents checked glottalization, double tick marks (′′) represent interrupted glottalization, and the tilde (~) represents palatization.

	Group	Description
-i′k	I	semantic extension
-ta•y	I	all
-muHk	I	bringing together
-tu•t	I	freeing
-ɨ	II	reciprocal, corporal (emotional or physical condition), goal-oriented
-ɨp	II	continued action
-ko	II	immediacy
-nɨ	II	already, with relative permanence

Table 2.3: Mixe Verbal Suffixes

Unlike verb roots, nominal, adjectival, and adverbial stems can have up to three syllables (with seemingly no internal structure). While nouns can sometimes created from verb roots via zero-derivation, others require slight modification. Often, for variable verb roots, the variant form is required for nominalization. In addition, there is a small set of nominalizing suffixes which grant a more specific meaning to derived noun. In general, there is no way to predict which method is necessary for a given verb [7: pp. 40–42].

Compound nouns consist of a modifier slot and a head slot [7: p. 42]. The former can contain any non-verbal stem, although compounds with adjectival or adverbial modifiers are relatively rare. The head slot must contain a noun stem.

There are two clitics in Mixe which are worth noting [7: pp. 43-44]. The postclitic *ɨk* is called the quotative marker, and is used to mark reported speech. It preferentially attaches to the first word of a clause. The postclitic *nɨ* carries a sense similar to the English word "still" and denotes incompleteness. It is often attached to "no" (*ka′*) yielding an expression (*ka′n*) which means "not yet."

Like verbs, nouns, adjectives, and adverbs can also take a variety of derivational affixes. Since all behave in a similar manner, prefixes and suffixes for each of the three nonverbal categories of stems will be treated together. Table 2.4 lists some of the more interesting examples [7: pp. 44–46]. Of these, one prefix, *ni•-*, merits specific attention. When combined with numbers, for example *mec* (two), the resulting word, in this case *ni•mec*, has a meaning of "[some number] together." This can be thought of as a generalization of the English word "both." Expressions equivalent to the English words *once*, *twice*, etc. are formed by the concatenation of the word for a number followed by the word *ʔo•k* [7: p. 50].

Mixe displays a few unique behaviors in its pronomial and locational deixis systems. Somewhat unusually, Mixe allows any number, not just *tu′′g* (one), to be used pronomially [7: p. 44][9]. Concerning deixis, apart from the relatively common

[9]The grammar (van Haitsma et al. [7]) is not exactly clear on the semantics, but presumably the usage

		Category	Description
ha•-		Noun	more, next (e.g. "next year")
pa-		Noun	wild (e.g. coyote/wolf vs. dog)
ʔa•-		Noun	related to the mouth, iterative (by analogy)
	-muk̆	Noun	curled together
	-ta′′g̃	Noun	woven
ko•-		Adj.	some amount of days hence
ko′′-		Adj.	some amount of days ago
mɨ•-		Adj.	relative order (e.g. second)
nɨ′-		Adj.	similar to English *-ish* concerning colors
ni•-		Adj.	together
ʔa•-		Adj.	iterative, together (with numbers)
	-paHk	Adj.	only (e.g. "only two")

Table 2.4: Mixe Nominal and Adjectival Affixes

three-way proximal-distal distinction[10], Mixe also has two demonstratives used to denote objects around, about, inside, or behind, the second of which connotes some sort of identification with the speaker [7: p. 49]. Moreover, for each demonstrative, there is a "specific" and "non-specific" version. The latter is used in questions inquiring about position or to express uncertainty.

Finally, I give a brief overview of Mixe questions. Unlike in English, the Mixe affirmative particle, *pa•y hadu′n*, always implies agreement with the statement of a yes-no question [7: p. 51]. Non-binary questions are answered by repeating the relevant clause, or by a word that is apropos to that which the question is inquiring about. Question themselves are formed in a couple of different ways [7: p. 102]. Interrogative questions use the word *neH*, and are maked by the particle *-(h/y)ɨ*. Rising intonation occurs on this particle. Echo questions are marked with the particle *ka′ hɨ*.

of other numbers as pronouns would refer to a group composed of however many indivudals (e.g. "those two," "these three").

[10] i.e. near the speaker, near the addressee, away from both the speaker and addressee.

2.8 South American

Guarani

Language Facts

Name	Guarani
Native Name	*Avañe'ẽ*
Language Family	Tupian
No. Speakers	2,500,000
Area Spoken	Paraguay

Dunsish's vowel harmony system was directly inspired by that of Guarani.

2.9 Sino-Tibetan

Bantawa

Language Facts

Name	Bantawa
Other Name(s)	Bantaba
Language Family	Sino-Tibetan, Kiranti
No. Speakers	166,600
Area Spoken	Nepal

2.10 Paleosiberian

Nivkh

Language Facts

Name	Nivkh
Other Name(s)	Gilyaki
Native Name	*N'ivxgu*
Language Family	Isolate
No. Speakers	1052
Area Spoken	Sakhalin

2.11 Constructed Languages

aUI

Conlang Facts	
Name	aUI
Creator	W. John Weilgart, PhD
Type	A priori
Category	Oligosynthetic Engelang
Year Created	1962

AllNoun

Conlang Facts	
Name	AllNoun
Creator	Tom Breton
Type	A priori
Category	Engelang
Year Created	1995

Davin

Conlang Facts	
Name	Davin
Creator	Zachary Weaver
Type	A priori
Category	Loglang
Year Created	*ca.* 2013

Davin is a set-theoretic loglang designed by Zachary Weaver [36]. Davin is refreshing in that it does suffer from the extreme naïveté that befalls most self-styled "logical" languages. On the contrary, its design exhibits its creator's competence in formal logic and set theory.

Davin is isolating, head-final, and ergative[11]. There are two major classes of words: *aŋtah* and *owpys*. *Aŋtah* are deictic words and proforms, as well as proper nouns, and always begin with a consonant. *Owpys* are non-proper nouns, verbs, adjectives, and adpositions, or in Davin parlance, "operators," and conversely always begin with a

[11]Due to the unfortunately limited amount of information on Davin available, it is difficult to determine what this precisely means. The best I can surmise is that all unary relations, the second operand of binary relations are syntactically identical.

vowel.[12] Weaver describes *aŋtah* as being *pushed*, and *owpys* as being *pushed then popped*. Presumably, this alludes to something akin to stack machine semantics, as Davin is syntactically similar to reverse Polish notation.

In any case, it appears that *aŋtah* essentially function as variables of a sort. They are also described as "pre-existing sets." However, *aŋtah* also function as quantifiers. *Owpys* are basically predicates, they denote *IS-A* relationships (e.g. "this is a ball," "that is a cube," etc.). *Owpys* application is left-associative. *Owpys* are by default intransitive (i.e. unary) but can be made transitive (i.e. binary) by infixing a nasal consonant after the initial vowel, which in turn assimilates to the place of articulation of the consonant which it precedes. Evidently, all *owpys* can be transitivized, but it appears that the semantics are largely unpredictable, and peculiar to each predicate. Included among *owpys* are the logical connectives. Missing parameters can be "threaded" through discursive context, filling in the proverbial blanks.

There are furthermore ten so-called "grammar words" called *aʃov*[13]. Several of these *aʃov* can switch the order of parameters, in a manner reminiscent of combinatory logic. *Aʃov* can also be used to form meta-linguistic utterances. For single *owpys*, the prepositional particle *eb* has a sense similar to the English suffix *-ness*. For phrases, the circumposition *yp lab…up* achieves the same effect. The particle *ejʒ* is used to parameterize relations, and thus form adverbial constructions. The particle *ow* is the quotation operator, and forms a powerful means of deriving new terms. Quoted words and/or phrases can be used much like any other.

Kalaba-X

Conlang Facts	
Name	Kalaba-X
Creator	Kenneth L. Pike
Type	A posteriori (vocabulary)
Category	Teaching Language
Year Created	1957

Kalaba-X is a pedagogical artificial language invented by Kenneth L. Pike in order to aid in teaching general translation skills [22]. While Pike presumably intended to demonstrate translation across *natural* languages, Kalaba-X nevertheless has several aspects which are more reminiscent of loglangs such as Lojban. Kalaba-X is perhaps unusual among conlangs in being designed by a professional linguist[14]. In truth, while Kalaba-X is described as a language, this is not entirely accurate. In his examples, Pike uses English words to illustrate the grammar. It is probably more correct to say that Kalaba-X is a skeleton of a language, a syntax and a grammar, lacking any concrete lexical representation.

[12] This is somewhat the opposite of Dunsish.

[13] I assume these are all particles of some sort.

[14] Though by no means unique. Tolkien, the ur-conlanger, was after all a highly accomplished philologist [5: p. 131].

According to Pike, each Kalaba-X sentences consist of three "slots," predicate, object, and subject. Each slot is mandatorily filled, and always occurs in the stated order[15]. Consequently, there is no distinction made between transitive and intransitive verbs (or in Kalaba-X parlance, *predicates*): all verbs are transitive. Pike calls this three item template a "formula," and further states that all sentences are built according to the singular form given. In addition, there is an optional "modifier" slot after each of the three mandatory slots. Unlike English adjectives, Kalaba-X modifier phrases are head-initial (like Latin or French). Figure 2.4 shows the complete so-called "sentence formula." In Pike's original notation (shown in 2.4a) the plus sign (+) is used to denote a mandatory slot, while the plus-minus sign (±) refers to an optional slot. The forward slash (/) represents a choice between options (cf. the "pipe" character in regular expressions). I have also included my own notation (shown in 2.4b). Square brackets represent optionality and the pipe unordered choice.

(a) $+(P^{V} \pm M^{N/V/A}) +(O^{N} \pm M^{N/V/A}) +(S^{N} \pm M^{N/V/A})$

(b) verb [verb|noun|adj] noun [verb|noun|adj] noun [verb|noun|adj]

Figure 2.4: Kalaba-X Sentence Formula

Pike never published a grammar or extended description of Kalaba-X, and to my knowledge, the only written record of it is contained in the present article [22][16]. Nevertheless, Pike gives numerous examples of translations into Kalaba-X, so the structure of the language can be surmised even from the relatively meager corpus. Curiously, there are no pronouns in Kalaba-X. Rather, constructions such as "speaker present" (i.e. the present speaker) are used instead. Similarly, there is no inflectional category indicating possession (or indeed, any inflection at all). Rather, a Kalaba-X speaker would merely use the predicate "own" (cf. *The bike which I own* vs. *My bike*). Kalaba-X, lacking any sort of particle or adposition, expresses non-core grammatical relations lexically. As mentioned all three slots are *always* mandatory. Thus, in order to express what in English one might use the passive voice for, some sort of "dummy" subject must be used (cf. English *it rains*). In the case of zero-object sentences, seemingly instransitive predicates can instead be easily reworked into reflexives. An interesting pattern occurs when nouns or verbs are used as modifiers. A syntax diagram of sentences 2a and 2b are given in figure 2.5.

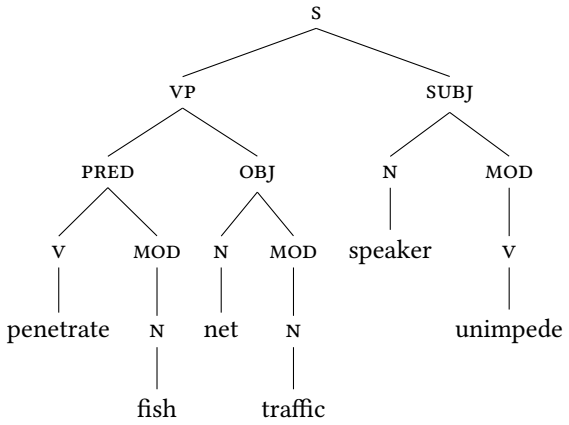
(1) Being a very small fish, I had got through the net. *English*

(2) a. Constituted analogy fish small speaker. *Kalaba-X*

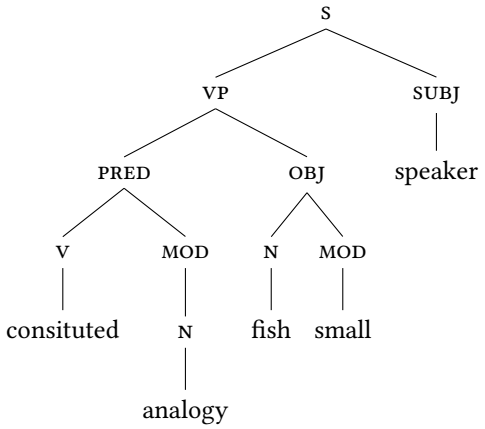
b. Penetrated fish net traffic speaker unimpeded.

[15]Essentially, this order of predicate, object, and then subject corresponds with the typologically unusual VOS word order. Perhaps Pike's intent in choosing such an unusual basic word order was to underscore the significant syntactic and grammatical differences between the language and English, inkeeping with his stated goal of teaching translation.

[16]Which itself was published in the journal of a seminary, making Kalaba-X the only overlap, at least of which I am aware, between conlanging and academic theology.



(a) Sentence 2a



(b) Sentence 2b

Figure 2.5: Kalaba-X Syntax Trees

In rendering the English sentence 1 into Kalaba-X, Pike treats the metaphorical usage of the word "fish" adverbally. Kalaba-X does not appear to allow for any use of recursion; complex discourse must instead be broken into and threaded through simple sentences. This is perhaps an artefact of pre-Chomskyan linguistics.

Okuna

Conlang Facts	
Name	Okuna
Creator	Matt Pearson
Type	A priori
Category	Artlang
Year Created	1993

Old Albic

Conlang Facts	
Name	Old Albic
Native Name	*Elbirin*
Creator	Jörg Rhiemeier
Type	A priori (?)
Category	Artlang
Year Created	2001

Old Albic (hereafter OA) is an absolutely beautiful artlang that easily ranks among my most loved conlangs [27]. Unlike Dunsish, OA has an internal history. However, unlike many artlangs, OA is not set on a fictional world, but rather is posited as a hypothetical "Old European" language, which predated the arrival of the Celtic, Indo-European languages on the British Isles. While I mostly haven't *directly* borrowed any features from OA, in a more general sense it has served a source of immense aesthetic inspiration.

With that in mind, there are a number of similarities between Dunsish and OA which are largely due to common influences (Celtic languages, etc.). Both languages have VSO as their primary or neutral word order[17]. Both languages make use of a hybrid or intermediary system of case-marking inbetween agglutinative and IE-style fusion. Both languages feature *suffixaufnahme*. Finally, both languages have somewhat similar structural limitations on root morphemes.

Phonologically, OA has an unusually plausible and well thought-out system[18]. There are seven vowel phonemes which feature a phonemic binary length contrast. However, "long" vowels differ from their short counterparts additionally in tenseness. OA's system of vowel harmony can be thought of as an eight vowel phoneme, represented by /○/, which absorbs the features of whichever vowel are nearest to it. In addition to vowel harmony, OA undergoes a form of assimilation termed umlaut[19].

[17] In-universe, the presence of VSO word order in the insular Celtic languages is explained as a product of an OA substrate.

[18] At least, in comparison to many conlangs where the "phonology" is little more than a set of phonemes.

[19] similar to the umlaut of Germanic languages

Umlaut is triggered by the presence of /a/, /i/, or /u/ and procedes from right to left, even across root boundaries. OA has a non-phonemic pitch accent, which is predictable based on word structure. On long vowels, two kinds of pitch are possible, which Rhiemeier terms *thrusting* and *slipping*. Thrusting tone has one peak, while slipping tone has two. Finally, words undergo a process of external sandhi called *linking* which is similar to French *liaison*.

Both lexical roots and affixes have restrictions on their shape. Roots are generally of the form CVC, CRVC, or CVRC, where 'R' represents the phonemes /m/, /n/, /ng/, /l/, /r/, /v/, or /j/ (i.e. the resonants). While some roots are bisyllabic, in such cases the vowels of both syllables are identical. Affixes are comparitively simple. Suffixes may be of the forms -C, -V, -VC, or -CV (-CVC does occur, however). Prefixes are generally CV-, but may be C- or V-. Most affixes have /○/ as their vowel, though /i/ and /u/ also occur.

Words in OA are composed of at least one root plus optional derivational affixes and any necessary inflectional morphemes. Derivation is mostly accomplished via suffixation, but there are a few prefixes. Prefixes never change the part of speech of a root, while suffixes can and often do. Prefixes may, however, change verb valency. Derivational affixes always come before inflectional affixes. Nouns may occasionally appear as a bare root, while verbs always have at least one inflectional affix attached. Compounds are common in OA, and are usually head-final. Part of speech, absent any derivational suffixes, is determined by the second component. Noun compounds are the most common, and of these most are attributive.

OA's gender system has two tiers. On the top level is an animacy distinction. Gender assignment is mostly semantic, with a few idiomatic exceptions. Morphologically, animate nouns can take a greater amount of cases than inanimate nouns. Among animate nouns there is a further subdivision into masculine, feminine, and common or neuter. Again, gender is assigned in a largely semantic manner. In situations where gender is unknown, irrelevant, or inapplicable, common gender is used. Gender is marked by final vowel: *-o* for masculine, *-e* for feminine, and *-a* for neuter. OA has three numbers: single, dual, and plural. Dual is marginal and not productive. It is only used for natural pairs (e.g. eyes) and is treated identically to the plural with respect to grammatical agreement.

OA has a total of ten different cases: agentive, genitive, dative, partitive, objective, instrumental, locative, allative, ablative, and perlative. Of these, only animate nouns may inflect for the agentive, genitive, dative, or partitive cases. The first four case endings are attached to the agentive stem, whereas the rest are attached to the objective stem. Only animate nouns have agentive stems, and as such inanimate nouns are precluded from improper inflection. The case are mostly self-explanatory, but there are a few exceptions which require further explanation. The partitive case, apart from representing part-whole relationships, is also used to mark inalienable posession. Conversely, the genitive is used only for alienable posession. The instrumental functions as a comitative when used with animate nouns. Moreover, the instrumental is also used to form adverbs and adjectives. The locative can sometimes express partitive relationships or inalienable posessesion with inanimate nouns. Nouns of persons have an extra *possessive local* case, formed by joining the locative case suffix to the end of a genitive inflected noun, which refers to someone's place (e.g. home). The dative,

partitive, and the locative are used with prepositions. The genitive, partitive instrumental, and locative undergo *suffixaufnahme*[20], though for the latter three only when used to express possession. Suffixaufnahme is the phenomenon by which possessors behave like adjectives, agreeing with the possessum in case, number, and gender. The possessum itself is in the *construct state*: it is definite without the presence of an article and can appear with its bare objective stem, case, number, and gender being expressed on the possessor.

Adjectives agree with the head in gender/animacy, case, and number. Adjectives can also be marked for comparison. The comparitive and superlative degrees are marked with suffixes. The equative degree (i.e. as ... as ...) is marked by a suffix and the object of comparison is then placed in the locative case. Adjectives function as adverbs when in the instrumental. The definite article similarly agrees with the noun in the same three categories and is placed before its head.

OA does not have true prepositions. Generally, noun cases are sufficient to express such relations. However, in such cases where more specificity is required, nouns expressing local relations are used to accomplish the same effect. The relation noun is placed in a locative case, and the noun of which the relation is over is put in the appropriate possesive case.

Pronouns inflect for all three numbers and case. First person pronouns feature a clusivity distinction. Second person pronouns distinguish between a familiar and honorific form. Third person pronouns inflect for animacy and gender. All three persons have separate so-called *emphatic* and reflexive forms. Moreover, OA has a set of *switch-reference* pronouns, *ra* and *ram*, which refer to the patient and agent of the preceding clause, respectively. Pronouns of core arguments are usually dropped. Demonstratives have a three-way proximal-distal distinction.

OA verbs agree with both the subject and the object (when applicable). Verbs themselves come in three varieties: active, stative, and fluid. Active verbs are those which refer to actions performed by a subject. Transitive verbs are always active. Stative verbs are the rest. Fluid verbs can function as either active or stative depending on whether the action is performed by the subject. Subjects of active verbs are marked either in the agentive or dative cases, depending on the volitionality of the action, and thus must be animate. Inanimate "subjects" are marked for the the instrumental case, and the verb is given a zero-subject. Subjects and direct objects of stative verbs are both marked for the objective case. Strong verbs are those which are made up of basic verbal roots, whereas weak verbs are derived from nouns, adjectives, etc. The two differ in inflections, with the former generally being more irregular. Prefixes called *preverbs* can promote oblique argument to direct object, and demote a direct object to the instrumental. Verbs also inflect for *version*. Neutral version is default and unmarked. Centripetal and centrifugal version indicate that the subject acts for herself or for others, respectively.

As mentioned, the basic word order in OA is VSO. Heads generally come before dependants. However, normal or neutral word order can be overridden to topicalize a particular word. Instead of a passive voice, OA allows for zero-subjects (in which case the verbs does not inflect for subject). The choice of the agentive or dative case for the

[20]In German, literally "suffix-absorption"

subject of stative verbs determines the degree of volition. For perception verbs, use of the subjective implies a deliberate perceptive act (e.g. looking vs. seeing). Subjects can also inflect for the instrumental, implying that the action is being caused by some other force. Instead of a copula, OA simply inflects the objective stem of a noun (marked for gender to agree with the subject) like a weak stative verb.

Sona

Conlang Facts

Name	Sona
Creator	Kenneth Searight
Type	A priori
Category	Oligosynthetic Auxlang
Year Created	1935

Sona is an early attempt at an oligosynthetic auxlang [32]. It consists of 360 roots, termed *radicals* by Searight, that are used to derive all words in the language [8]. Table 2.7 lists the 360 Sona radicals described in Searight's book [32: ch. 8]. Like many such languages created during the early 20th century, Sona shows some signs of a naïve relexification of English. For example, Sona tense morphology (though not the morphemes themselves) is nearly identical to that of English. The main challenge that all oligosynthetic languages face is dealing with proper nouns which do not cleanly fit into the often rigid morphophonological constraints that such languages inevitably have. Interestingly, Sona seems to have no such "escape hatch," save for orthographically: proper nouns are capitalized.

Sona's phonology is well-designed, with a simplicity befitting an auxiliary language (as opposed to Esperanto, for instance). Sona has a typical five vowel system, /a e i o u/, each of which is pronounced essentially in the same manner as Spanish. The consonants /g k d t z s m n b p l v f h/ have the same values as in IPA. The letter /r/ may be prounounced as any rhotic consonant. The letters /c/ and /j/ represent the affricates [t͡ʃ] and [d͡ʒ] respectively, while /x/ represents the palatal fricative [ʃ]. With the exception of /m/ and /n/, /l/ and /r/, and /h/ and /x/, radicals are arranged in voiced–voiceless pairs. With respect to phonological structure, radicals may divided into two groups by their root consonants: stops, nasals, and the sibilants /s/ and /z/; and the affricates and fricatives (save the aforementioned). Radicals of the former group always have 25 of each. Of these 25, they are divided further into five subgroups of five radicals each. Each subgroup differs by having a different "thematic" vowel. There are five different forms: CV, CVn, aCV, iCV, uCV. In this notation, uppercase "C" represents the root consonant while uppercase "V" represents the thematic vowel of the subgroup. Lowercase letters simply denote the given phoneme. Conversely, the latter group of radicals have only 10 each, of which there are five pairs, differing in thematic vowels, of the forms CV and CVn. It is not clear why Searight chose to assign only 10 radicals each among the members of this set. Harrison refers to radicals of the form CV as "primary" and those of the forms CVn, aCV, iCV, and uCV as "secondary."

As mentioned, Sona's phonological system is in general a marked improvment over other proposed auxiliary languages, especially its contemporaries. Sona's phoneme inventory is nearly the same as the 15 consonant and 5 vowel system proposed by Jörg Rhiemeier [26]. However with respect to Searight's stated goal of producing an international auxiliary language, there are a number of problems. In particular, the fact that voiced–voiceless stops and the liquids /l/ and /r/ can form minimal pairs would prove challenging to speakers of many languages. Moreover, while most languages (numerically) have a five vowel system, there are several languages with large amounts of speakers, Arabic chief among them, which only have the three vowel system /a i u/[21]. Thus, as with consonants, an auxiliary language should minimize the functional load between the pairs /e/ and /i/, /o/ and /u/, /a/ and /e/, and /a/ and /o/. Another strength of Sona's design is that while CVC syllables do occur, only the nasal consonant /n/ may occur in the coda. This is a reasonable compromise between only CV syllables, which is too restrictive, or the full set of CVC syllables which would be complicated for many speakers. A possible extension would be to allow roots with /m/ or /r/ as the coda.

There does not appear to be any systematic relationship between pairs of roots differing in the voicing of their root consonant (or between the pairs /m/ and /n/, /l/ and /r/, or /h/ and /x/). This is probably for the best. If there are to be minimal pairs between difficult to distinguish sounds, they should at the very least be of disparate semantic classes, to prevent confusion over what would be functionally treated by many as homonyms. Among the larger groupings of radicals sharing a root consonant, there likewise doesn't appear to be any consistent semantic relationship. Conversely, the subgroups of five or two members generally feature a vague semantic relationship with one another. Often, the relationship is metaphorical in nature. For example, the radicals of the RA group are associated with both masculinity and power. However, this is not always the case. For instance, the radicals *xi* and *xin* have the unrelated meanings "six" and "glaze, sheen" respectively. Some radicals have a rather eccentric collection of meanings: *bo* can mean flesh, blood, or German[22].

There are two sets of radicals which require further discussion. The first is the so-called "particles," listed in table 2.5. These as always begin with a vowel. The second are the 36 "indicators," some of which are also particles: *a*, *ba*, *bi*, *ci*, *da*, *di*, *e*, *en*, *fi*, *ga*, *gi*, *ha*, *he*, *i*, *in*, *ji*, *ka*, *ke*, *ki*, *ko*, *li*, *na*, *ne*, *ni*, *o*, *pan*, *po*, *ra*, *ri*, *ru*, *si*, *ta*, *to*, *u*, *un*, and *zan*. Indicators are always monosyllabic and, with two exceptions, always begin or end with a vowel. The difference between the two sets (and indeed, as mentioned, there are radicals common to both) is not clear. In any case, both groups have specific grammatical functions, beyond merely their semantic function. Notably among auxiliary languages, Sona has an extensive case system which is detailed in table 2.6 [32: §31]. While Searight refers to them as "cases," this terminology is arguably inappropriate. Of the 12 "cases," three are actually word order/syntax, and six are prepositions.

[21] Quechuan languages have the same three vowel system /a i u/. In many dialects of Peruvian Spanish, this causes a distinctive accent, called *motosidad*, stemming from the difficulty L1 Quechua speakers have in correctly pronouncing Spanish's five vowel system [15].

[22] Perhaps not too surprising considering the era in which Sona was created!

Where Sona displays the most originality is in its system of compounding and derivation. Interestingly, radicals have different meanings when used as prefixes and suffixes. The radical *na* as a prefix means *non-*, while as a suffix is basically used to denote the sense of an inorganic substance or object. Sona compounds are always head-final [32: §21–22]. Searight, following the "Arab Grammarian[s]," classifies all radicals into one of three categories: noun, particle, or verb [32: §6]. This is similar in spirit to Saussere's threefold division of Esperanto roots (see section 5.1). Searight generally prefered terse coinings, and appears to have accordingly viewed radicals, like Latejami, as primarily mnemonic components of self-standing words. Which is to say that the intent is not to perfectly categorize all concepts into a hierarchy, or to allow a speaker of the language to be able to precisely derive the meaning of a word from those of its components. Rather, the radicals are loose, often vague, classifiers which are meant to give clues or hints as to meanings of words in which they are compounded. The consequence of this is that Sona compounds would be expected to take on lexicalized uses, with definitions which extend beyond the mere combined meaning of the radicals.

One difficulty is that morpheme boundaries, unlike X-1, are highly ambiguous. That being said, Sona does have a few features which help to ameliorate this. The letter /y/ (presumably [j]) is used to disambiguate morpheme boundaries. It is inserted between syllable-final /n/ and a subsequent vowel and between two adjacent identical vowels, with the exception of the sequence /ii/. It is not clear how Searight dealt with ambiguities in this particular instance. Finally, /y/ is also inserted between adjacent "hard" vowels, /a e o/[23]. The syllable /ci/ can be used to clarify compounds which are still ambiguous. The sequence /ye/ is inserted between a vocalic radical prefixed to radicals beginning with one of /g k d t z s m n b p l r/. In certain instances, vowels may be elided. If a secondary radical is followed by a corresponding primary radical, then the common thematic vowel is dropped. Similarly, if a aCV, iCV, or uCV is followed by a CVn radical of the same group, then the thematic vowel is also likewise dropped. The implication is that consonant gemination is thus phonemic.

Table 2.7: Sona Radicals

G				K		
	GA	*n.*	physical, organic	**KA**	*v.*	lead, cause, chief
	gan	*n.*	matter, substance	**kan**	*n.*	house, build, construct
	aga	*n.*	earth, ground, land	**aka**	*p.*	above, high, up, raise
	iga	*n.*	salt, alkalai, soda	**ika**	*n.*	point, sharp, prick, p. just
	uga	*n.*	grain, seed	**uka**	*n.*	vertical, stand up
	GE	*p.*	and, join, chain, link	**KE**	*p.*	what?, query, ask
	gen	*n.*	compound, mix, blend	**ken**	*p.*	if, doubt, suspect
	age	*v.*	meet, contact	**ake**	*v.*	request, propose, offer
	ige	*v.*	cohere, stick, cling	**ike**	*v.*	suggest, guess, hint, suppose
	uge	*n.*	cumulus, pile, heap	**uke**	*v.*	seek, hunt, quest, search

Continued on next page

[23]What I gather is that "hard" refers to the fact that pairs of these vowels form diphthongs that are uncommon or difficult for many individuals to pronounce.

Table 2.7: Sona Radicals — continued from previous page

GI	*n.*	collect, gather, group	**KI**	*n.*	start, begin, ready, prepare
gin	*n.*	texture, weave, net, web	**kin**	*n.*	origin, root, cause
agi	*n.*	compress, dense, thick	**aki**	*n.*	speed, quick, hasten
igi	*v.*	mix, confuse, puzzle	**iki**	*n.*	sudden, shock, flash
ugi	*n.*	tangle, bush, cluster	**uki**	*v.*	develop, train, culture
GO	*n.*	circle, round, wind, roll	**KO**	*n.*	small, few, less, child
gon	*n.*	sphere, ball, round	**kon**	*n.*	minute, atom, germ, speck
ago	*n.*	curve, bend, arc	**ako**	*n.*	short, brief, clip, shrink
igo	*n.*	convex, round, cheek	**iko**	*n.*	narrow, thin, squeeze
ugo	*v.*	fold, bind, wrap, clasp	**uko**	*n.*	frail, slender, delicate, fine
GU	*n.*	mouth, gap, gulf, gulp	**KU**	*v.*	lie, supine, bed, lay
gun	*n.*	angle, corner, fork	**kun**	*n.*	base, seat, sit
agu	*n.*	slope, lean, tilt, trend	**aku**	*n.*	surface, skim, slip, slide
igu	*n.*	cone, horn, wedge	**iku**	*n.*	flat, level, horizon
ugu	*n.*	hook, claw, catch	**uku**	*n.*	layer, shave, slice, film
DA	*v.*	do, act, p. to (Infin.)	**TA**	*n.*	great, augment, p. very
dan	*n.*	value, money, price	**tan**	*n.*	body, bulk, mass, 1000
ada	*n.*	work, function, make	**ata**	*v.*	spread, stretch, wide
ida	*v.*	direct, office, minister	**ita**	*v.*	fill, charge, thorough
uda	*n.*	law, justice, code	**uta**	*v.*	swell, wave, undulate
DE	*p.*	through, hole, pierce	**TE**	*n.*	hand, project, take
den	*n.*	tooth, bite, plough, dig	**ten**	*v.*	hold, contain
ade	*n.*	hollow, cave, cup, scoop	**ate**	*n.*	tube, cylinder
ide	*n.*	notch, dent, crack	**ite**	*n.*	flap, leaf, blade, tongue
ude	*v.*	break, rend, burst	**ute**	*n.*	pocket, sheath
DI	*p.*	of, part, limb	**TI**	*n.*	they, them, their
din	*n.*	party, sect, creed	**tin**	*n.*	three, triangular
adi	*n.*	portion, share, lot	**ati**	*v.*	throw, shoot, missile
idi	*v.*	scatter, sow, spray	**iti**	*n.*	plan, object, intend
udi	*v.*	mangle, wound, tear	**uti**	*p.*	beyond, far, distance
DO	*n.*	two, both, bi-	**TO**	*v.*	pass, year, was, were
don	*n.*	give, grant, issue	**ton**	*v.*	tire, wear, use up
ado	*v.*	add, affix, sum	**ato**	*n.*	old, age, senior
ido	*v.*	advise, prompt, urge	**ito**	*n.*	end, final, expire
udo	*v.*	pay, reward, atone	**uto**	*n.*	excess, flood, p. too
DU	*n.*	wood, nut, hard	**TU**	*n.*	you, your
dun	*n.*	horn, shell, crust	**tun**	*n.*	limit, wall, city
adu	*n.*	bone, thin, skeleton	**atu**	*n.*	eight, octagonal
idu	*n.*	brittle, snap, rotten	**itu**	*n.*	cube, block, dice
udu	*n.*	thump, thud, drum	**utu**	*v.*	swell, tumour, lump

D T

Continued on next page

Table 2.7: Sona Radicals — continued from previous page

Z	**ZA**	*p.*	back, rear, ago	**SA**	*v.* know, teach, inform
	zan	*n.*	woman, female	**san**	*n.* health, medicine, sane
	aza	*v.*	delay, slow, late, lazy	**asa**	*v.* prove, certain, ensure
	iza	*n.*	pause, hover, hesitate	**isa**	*n.* skill, talent, craft, art
	uza	*v.*	yield, give up, submit	**usa**	*v.* heed, obey, comply
	ZE	*n.*	evil, sin, ugly	**SE**	*v.* see, seem, visual
	zen	*n.*	ill, sick, disease	**sen**	*v.* feel, touch, sense
	aze	*v.*	acuse, blame	**ase**	*v.* care, watch, keep
	ize	*n.*	clumsy, stupid, blunder	**ise**	*v.* find, perceive, solve
	uze	*n.*	harm, hurt, corrupt	**use**	*v.* hear, sound, listen
	ZI	*n.*	real, fact, positive, be	**SI**	*n.* self, one, own, identity
	zin	*v.*	live, life, vital	**sin**	*n.* form, shape, type
	azi	*n.*	vigour, active, vivid	**asi**	*n.* like, same, copy
	izi	*n.*	ease, simple, light	**isi**	*n.* mutual, exchange
	uzi	*v.*	stir, wake, inspire	**usi**	*v.* agree, fit, conform
	ZO	*n.*	fire, burn, light	**SO**	*p.* with, help, combine
	zon	*n.*	poison, infect	**son**	*n.* hundred, per cent
	azo	*n.*	acid, bitter, sarcasm	**aso**	*n.* truth, faith, trust
	izo	*n.*	pain, hurt, ache	**iso**	*n.* equal, even, balance
	uzo	*n.*	spasm, cramp, writhe	**uso**	*n.* treaty, bargain, pact
	ZU	*n.*	animal, brutal, fierce	**SU**	*n.* water, liquid, flow
	zun	*n.*	seven, heptarchy	**sun**	*n.* fish, clammy
	azu	*n.*	storm, fury, violent	**asu**	*n.* slime, bog, mucus
	izu	*n.*	rough, crude, raw	**isu**	*n.* cold, chill, frigid, keen
	uzu	*n.*	wild, random, chaos	**usu**	*n.* silver, glint, steel
M	**MA**	*n.*	place, put, local	**NA**	*p.* not, no, nor, deny
	man	*v.*	stray, wait, remain, rest	**nan**	*p.* although, ignore
	ama	*p.*	instead, proxy, vice-	**ana**	*p.* aganist, hostile, oppose
	ima	*v.*	have, belong, get	**ina**	*v.* cancel, destroy, wipe
	uma	*n.*	mother, nation, home	**una**	*n.* sole, only, isolate
	ME	*v.*	think, mind, idea	**NE**	*p.* from, leave, absent, rob
	men	*n.*	measure, size, month	**nen**	*p.* except, omit, skip
	ame	*v.*	regard, note, observe	**ane**	*v.* avoid, shun, flee
	ime	*n.*	belief, opinion, convince	**ine**	*v.* seep, latent, numb
	ume	*n.*	reason, rational, logic	**une**	*n.* reserve, shy, retire
	MI	*n.*	I, me, my	**NI**	*v.* suffer, undergo, able
	min	*p.*	in, middle, interior	**nin**	*n.* product, child, result
	ami	*n.*	mean, modify, normal	**ani**	*p.* under, fall, low, down
	imi	*p.*	among, between, hiatus	**ini**	*n.* quiet, peace, soothe
	umi	*n.*	centre, heart, focus	**uni**	*v.* depend, hang, hable

Continued on next page

Table 2.7: Sona Radicals — continued from previous page

	MO	*n.*	fruit, bulb, nipple	**NO**	*v.* carry, vehicle
	mon	*n.*	profit, gain, reap	**non**	*v.* invert, reverse
	amo	*n.*	climax, acme, summit	**ano**	*p.* across, span, transfer
	imo	*v.*	suck, pump, digest	**ino**	*n.* new, mend, convert
	umo	*n.*	breast, milk, suckle	**uno**	*n.* cross, check, squint
	MU	*n.*	hair, fur, wool, brown	**NU**	*v.* change, turn, or
	mun	*n.*	soft, pile, cushion	**nun**	*n.* nine
	amu	*n.*	pliant, supple, flex	**anu**	*v.* upset, revolt, disturb
	imu	*n.*	smooth, mild, gentle	**inu**	*n.* other, vary, differ
	umu	*n.*	bear, bore, grumpy	**unu**	*n.* contrast, judge
B	**BA**	*v.*	strike, military, weapon	P **PA**	*v.* eat, nourish, bread
	ban	*v.*	contest, fight, match	**pan**	*n.* all, complete, fill
	aba	*v.*	punish, penal, avenge	**apa**	*n.* supply, store, (furnish), equi
	iba	*n.*	stubborn, persist, try	**ipa**	*p.* enough, suffice, cofort
	uba	*v.*	forbid, deny, ban, tabu	**upa**	*v.* cram, stuff, choke, boil
	BE	*p.*	without, strip, poor, (undo)	**PE**	*n.* foot, leg, walk, kick
	ben	*n.*	clean, bathe, polish	**pen**	*n.* five, pentagon
	abe	*n.*	clear, pale, fair, flax	**ape**	*n.* reptile with legs, creep
	ibe	*v.*	spend, waste, drain	**ipe**	*n.* reptile, sinuous, crawl
	ube	*n.*	empty, void, vacate	**upe**	*v.* jump, spring, skip, hop
	BI	*p.*	by, use, tool	**PI**	*n.* bird, fly, aviation
	bin	*n.*	utensil	**pin**	*n.* feather, fringe, flimsy
	abi	*n.*	habit, mode, custom	**api**	*n.* insect with wings
	ibi	*n.*	trade, commerce	**ipi**	*n.* insect, vermin
	ubi	*v.*	serve, tend, attend	**upi**	*n.* buoyant, float, swim
	BO	*n.*	flesh, blood, german	**PO**	*p.* on, by, cover, dress
	bon	*n.*	plenty, wealth, bounty	**pon**	*v.* weigh, heavy, load
	abo	*n.*	antelope, buck	**apo**	*p.* on, impose, ride, tax
	ibo	*n.*	cattle, buffalo	**ipo**	*n.* overlay, veneer, coat
	ubo	*n.*	sheep, goat	**upo**	*n.* shield, defend, shelter
	BU	*n.*	smell, fume, smoke, kiss	**PU**	*n.* filth, foul, dirt, soil
	bun	*n.*	taste, savour	**pun**	*n.* decay, pus, rot, rust
	abu	*v.*	love, desire, dear	**apu**	*n.* obscene, lewd, desire
	ibu	*n.*	choice, pick, critic	**ipu**	*n.* disgust, loathe, shame
	ubu	*n.*	greed, covet, yearn, lust	**upu**	*n.* swine, pig, pork
L	**LA**	*v.*	say, tongue, voice	R **RA**	*n.* man, male
	lan	*n.*	sound, play, music	**ran**	*n.* king, rule, royal
	ala	*v.*	praise, hail, cheer	**ara**	*n.* master, tame, tyrant
	ila	*n.*	bell, ring, jangle	**ira**	*n.* strength, fort, muscle
	ula	*v.*	sing, recite, howl, choir	**ura**	*n.* force, drive, electric

Continued on next page

Table 2.7: Sona Radicals — continued from previous page

	LE	*v.* write, design, letter	**RE**	*n.* straight, line, long
	len	*n.* document, book, read	**ren**	*n.* cord, tense, thin
	ale	*n.* sign, mean, flag, crest	**are**	*n.* order, arrange, series
	ile	*n.* trace, mark, print	**ire**	*p.* again, recur, continue
	ule	*n.* stain, blot, spot	**ure**	*n.* right, regular, correct
	LI	*p.* to, direct, point	**RI**	*n.* time, while
	lin	*n.* edge, outline, profile	**rin**	*n.* rhythm, pulse, metre
	ali	*n.* side, wing, flank	**ari**	*v.* reach, gain, come
	ili	*p.* near, about, approx	**iri**	*n.* recoil, react, spring
	uli	*n.* close, proximate, intimate	**uri**	*n.* season, suit, opportune
	LO	*v.* open, free, undo	**RO**	*n.* colour, dye, paint
	lon	*n.* loan, let, lend, hire	**ron**	*n.* tone, tune, shade, mood
	alo	*v.* allow, leave, ticket	**aro**	*n.* beauty, adorn, fair
	ilo	*n.* loose, hang, relax	**iro**	*n.* light, shine, lamp
	ulo	*n.* oil, grease, flatter	**uro**	*n.* gold, glory, splendour
	LU	*v.* play, game, amuse	**RU**	*v.* go, move, mobile
	lun	*n.* absurd, trivial, vain	**run**	*n.* machine, apparatus
	alu	*v.* invite, host, guest	**aru**	*v.* draw, pull, charm
	ilu	*n.* jest, laugh, merry	**iru**	*v.* vibrate, shake, thrill
	ulu	*n.* monkey, mischief	**uru**	*v.* swing, sway, wag, flux
J	**JA**	*v.* rub, scratch, jerk, rough	C	**CA** *n.* four, square, rectangle
	jan	*v.* gnaw, grind, chew, rodent		**can** *n.* box, cell, case, room
	JE	*n.* stone, hard, stiff, stern		**CE** *v.* fix, secure, support
	jen	*n.* metal, courage, coin		**cen** *v.* stop, shut, bar, gate
	JI	*n.* person, human, private		**CI** *n.* which, refer, some
	jin	*n.* spirit, soul, ghost, fairy		**cin** *n.* grade, scale, step
	JO	*n.* God, idol, ideal, sacred		**CO** *v.* cut, kill, knife
	jon	*n.* magic, spell, occult		**con** *n.* crush, dust, powder
	JU	*n.* delight, happy, toy, glad		**CU** *v.* wish, want, hope
	jun	*n.* young, fresh, junior		**cun** *n.* whim, caprice
H	**HA**	*n.* name, call, address, O	X	**XA** *n.* merit, good, nice
	han	*n.* title, honour, noble		**xan** *n.* thanks, thank, gratis
	HE	*n.* any, general, public		**XE** *n.* feline, cat, sly
	hen	*n.* common, vulgar, cheap		**xen** *n.* canine, dog
	HI	*n.* class, category		**XI** *n.* six, hexagonal
	hin	*p.* after, follow, sequence		**xin** *n.* gloss, glaze, sheen
	HO	*n.* special, define		**XO** *v.* show, evident, witness
	hon	*n.* genuine, honest, rely		**xon** *n.* pomp, ceremony, flaunt

Continued on next page

Table 2.7: Sona Radicals — continued from previous page

	HU	*n.*	air, breath, wind, weather	**XU**	*n.* sweet, sugar, fragrant
	hun	*n.*	gas, vapour, melt	**xun**	*n.* guile, cunning, diplomacy
V	**VA**	*n.*	front, face, ahead	F **FA**	*n.* chance, luck, risk, perhaps
	van	*v.*	hide, mask, shade, night	**fan**	*n.* fancy, myth, pretend
	VE	*n.*	vegetable, herb green	**FE**	*n.* false, cheat, trap, spurious
	ven	*n.*	cereal, crops, corn	**fen**	*n.* defect, fault, wrong
	VI	*n.*	nature, quality, kind	**FI**	*v.* fail, miss, slip, hardly
	vin	*n.*	essence, cream, wine	**fin**	*n.* feeble, weak, faint
	VO	*v.*	grieve, weep, woe, sorry	**FO**	*v.* fear, alarm, danger
	von	*v.*	pity, spare, mercy	**fon**	*n.* loud, noise, roar
	VU	*n.*	vague, wanter, stray	**FU**	*p.* out, exterior, foreign
	vun	*v.*	twist, coil, curl, screw	**fun**	*n.* odd, strange, peculiar

High Valyrian

Conlang Facts

Name	High Valyrian
Native Name	*Valyrio udrir*
Creator	David J. Peterson
Type	A priori
Category	Artlang
Year Created	2012

High Valyrian is a language designed by David J. Peterson for the HBO TV series *Game of Thrones* [16]. In-universe, Valyrian functions in a role analogous to that of Latin in the European Middle Ages [35]. With respect to Dunsish, there are two features of Valyrian which are of particular interest: the gender system and number.

Valyrian's gender system is perhaps its most unique and interesting feature. There are four genders: solar, lunar, terrestrial, and aquatic, which in Valyrian are called *vēzenkon qogror*, *hūrenkon qogror*, *tegōñor qogror*, and *embōñor qogror*, respectively [18]. Valyrian gender has characteristics of both Indo-European type gender systems as well as Bantu noun classes. Gender is mostly predictable from phonology (cf. Spanish) but there are also loose semantic trends to each gender and some semi-regular derivational patterns when changing the gender of a root[24]. Most animate, and what Peterson terms "individuatable" nouns are either solar or lunar. Diurnal animals tend to be solar, while nocturnal animals tend to be lunar. Names of (human) occupations and external body parts are generally solar. Words for food, plants, and

[24]Unfortunately, I was not able to discern what exactly these patterns *are*, just that they apparently exist.

Table 2.5: Sona Particles

A	**a**	*p.*	at, fixed place (sf.)
	an	*n.*	she, her
	ua	*p.*	for, as for, c. that
E	**e**	*n.*	many, frequent, plural
	en	*n.*	it, its, the
	ue	*p.*	ought, duty, oblige
I	**i**	*p.*	while, when (sf.)
	in	*n.*	this, the latter
	u	*p.*	yes, assent
O	**o**	*p.*	honorific, quality (sf.)
	on	*n.*	he, him, his
	uo	*p.*	must, necessity
U	**u**	*p.*	how, manner (sf.)
	un	*n.*	that, the former
	au	*p.*	but, alternative

Table 2.6: Sona Cases

Nominative	$\emptyset$	
Vocative	**-ha**	O x!
Accusative	$\emptyset$	
Directive	**-li**	for x
Motive	**dili**	for x's sake
Dative	**li**	to x
Locative	**a-**	at x
Ablative	**ne**	from x
Instrumental	**bi**	by x
Genitive	**vi**	of x
Partitive	**di**	(composed) of x
Posessive	$\emptyset$	x's

metals are often terrestrial. Words for liquids and bodies of liquids are usually aquatic. Finally, words for military objects (e.g. swords, shields, etc.) are most often lunar. Valyrian has several declensions which interacts with the gender system in various ways, similar to Latin, and an interesting pattern occurs at their intersection. There is a tendency for words which are duals of one another, for example *tala* and *trēsy*, son and daughter, to have the same gender but belong to different declensions, in this case lunar gender, but first and second declension respectively [20]. Adjectives agree with their head nouns in gender, although some adjectives conflate lunar and solar and terrestrial and aquatic into just two forms. Valyrian was a direct inspiration for Dunsish's elemental gender system (see chapter 6).

The Valyrian system of grammatical number is interesting as well[25]. Valyrian has four numbers: singular, plural, paucal, collective [19]. While the singular and plural are self-explanatory, the former two require some explanation. The paucal refers to a small amount of something, similar to "some" or "a few" in English. The collective, on the other hand, has a range of meanings, but generally means a large group of something, a sort of natural grouping, or all such things. Words referring to humans, besides those referring to professions, generally have the former meaning. Collective forms of inherently female nouns (e.g mother, daughter) tend to have gender-inspecific semantics (e.g. parents, children). Often, collectives and paucals are lexicalized, and become reanalyzed as words in their own right. For example, *azantyr*, "army," was originally the collective of *azantys*, "soldier." Similarly, *tīkun*, "wing," is a reanalyzed paucal form of *tīkos*, "feather." Lexicalized collectives and paucals, together with foreign loanwords, form the sixth declension class of Valyrian's noun system. Both verbs and adjectives agrees with subjects and heads (respectively) in gender. However, with respect to agreement, the collective is treated as singular and the paucal as the plural [21].

Finally, a worth interesting features of the adjective system are worth noting. In addition the to comparative and superlative, Valyrian also has an equative degree of comparison. This is often used for apposition. Syntactically, Valyrian is generally head-final. However, head-initial constructions are also possible and have a variety of pragmatic effects. Adjectives placed before nouns take shortened inflectional endings. Unlike Latin, substantives cannot be formed via zero-derivation.

Ygyde

Conlang Facts

Name	Ygyde
Creator	Andrew Nowicki and Patrick Hassel-Zein
Type	A priori
Category	Oligosynthetic Auxlang
Year Created	*ca.* 2009

[25] Grammatical number (e.g. singular, plural), as opposed to a numeral system.

X-1

Conlang Facts	
Name	X-1
Full Name	*Experimental Language 1*
Creator	Jörg Rhiemeier
Type	A priori
Category	Loglang
Year Created	2005

X-1 is a logical language (or *loglang*) notable for its use of a technique called *self-segregating morphology* [28]. In essence, self-segregating morphology describes a system wherein the morphophonology of a language is such that morpheme boundaries (and sometimes word boundaries) are never ambiguous (cf. self-sychronizing codes in coding theory). X-1 adopts a strategy in implementing self-segregation that is basically an improved version of Prothero's *Plan B*(see [23]).

The phoneme inventory of X-1 is minimal: there are seven consonants /p t k s l m n/, and four vowels /i ɛ ɔ u/. All syllables are of the form CV or V. There are eight possible onsets, consisting of the seven consonants and zero. Vowels can be thought of as a double (±F, ±H), where *F* and *H* refer to frontness and height, respectively, yielding the four possible vowels. The X-1 alphabet is described in table 2.8. An examination of the alphabet reveals a few different patterns. First, the order of the eight possible onsets is mirrored. Second, for the first half of the alphabet, obstruents (i.e. plosives and fricatives) are assigned odd numbers while sonorants (including zero onset) are assigned even numbers; this is reversed for the second half. Finally, frontness (±F) is also determined by the final bit of the binary representation (though for both halves): front vowels (+F) are even, while back vowels (–F) are odd. The consequence of this is that two numbers which are ones' complements (i.e. bit flipped) of one another always have the same onset and opposite frontness. However, the nucleus vowel is underspecified by the onset. Height (±H) is determined by the most-significant bit (MSB) of the binary representation of the following symbol. If the following symbol has an MSB of 1, then the preceding symbol's vowel is low (–H). Conversely, if the following symbol has an MSB of 0, or the given symbol is final (i.e. there is nothing after it), then the vowel is high (+H).

As mentioned, X-1 features a self-segregating morphology. This is acheived by encoding the length of the morpheme (in syllables/symbols) into it's first syllable. In table 2.8, the number of syllables that a morpheme beginning with any which symbol can have is given in the right-most column. These lengths are not arbitrary. Rather, they are a function of the number of consecutive one bits starting from the left of the binary representation, plus one. Morphemes beginning with *h* are treated specially: the length is given by the symbol which follows, thus allowing for morphemes of unbounded length. Morpheme length itself also encodes part of speech. The morpheme *j* delimits variable scope boundaries, that is the context of variable binding. The rest of

№	Hex	Binary	Symbol	Onset	Vowel	Syllables
0	0	0000	j	∅	+F	1
1	1	0001	g	[k]	−F	1
2	2	0010	l	[l]	+F	1
3	3	0011	z	[s]	−F	1
4	4	0100	ñ	[n]	+F	1
5	5	0101	d	[t]	−F	1
6	6	0110	µ	[m]	+F	1
7	7	0111	b	[p]	−F	1
8	8	1000	p	[p]	+F	2
9	9	1001	m	[m]	−F	2
10	A	1010	t	[t]	+F	2
11	B	1011	n	[n]	−F	2
12	C	1100	s	[s]	+F	3
13	D	1101	r	[l]	−F	3
14	E	1110	k	[k]	+F	4
15	F	1111	h	∅	−F	5+

Table 2.8: X-1 Alphabet

the single-syllable morphemes are variables[26]. Two-syllable morphemes are so-called *connectives*[27]. Morphemes with three or more syllables are *predicates*, and are the only open class of words (both by design, and structural necessity). Within the class of predicates, arity is furthermore encoded by length, subtracted by two. Three-syllable predicates are unary, four-syllable predicates are binary, and so forth.

X-5

Conlang Facts

Name	X-5
Full Name	*Experimental Language 5*
Creator	Jörg Rhiemeier
Type	A priori
Category	Arithmographic Language
Year Created	2007

[26] Rhiemeier is not clear on the precise way in which variables function, but presumably they are some sort of generalized proform or deixis.

[27] Again, the documentation is not clear. Most likely, "connectives" refers to *logical connectives* (e.g. negation, conjunction, dijunction).

While there is very little published information on X-5, it is said to be an arithmographic language [29]. *Arithmographic* is a term coined by Jörg Rhiemeier to describe the sort of system, first envisaged by G. W. Leibniz, where semantic primes are encoded in prime numbers [25]. In essence, such a scheme works by first assigning each semantic prime, understood as an intenally non-decomposable concept or idea, to a prime number. To represent complex or compound ideas, the numbers reprenting primes are multiplied together, and the compound idea is encoded as the product of the corresponding numbers of each of its constituent primes. According to the Fundamental Theorem of Arithmetic, every number besides 1 is either prime, or has a unique prime factorization [38]. Therefore, every complex idea can be encoded uniquely as the product of two or more semantic primes. Rhiemeier for example suggests that the concept of *life* could be encoded as the prime number 2, thus every even number would refer to something relating life, and in turn every odd number the opposite.

The arithmographic language has a number of advantages over the more conventional taxonomic or hierachical variety of philosophical languages, typified by Wilkins' *An Essay Towards a Real Character and a Philosophical Language* [39]. First, the set of primitives is unbounded, given that there is an infinite number of prime numbers. As such, the system is less vulnerable to the rigidity imposed by a strict and finite hierarchical system of classification. Moreover, semantically close words do not have similar sounds. Rather, some sort of algorithm is used to bijectively map the set of natural numbers onto the set of phoneme strings.

Conversely, the primary weakness of the arithmographic language is the relationship between primes is ambiguous as no priority or role is assigned to any particular component. Generally in, for example, binary compounds, there is both a *head* and a *dependent*, of which the latter modifies the former, cf. the difference between "eyeglasses" and "glass eyes." Similarly, such a scheme is particularly vulnerable to the so-called "bracketing paradox" (see [34]). Consider the famous example of a "pretty little girls' school." Is this a pretty school, which is for little girls? Or is it a pretty and little school, for girls? Or perhaps a school only for pretty little girls?

One possible solution to this problem is to employ some more sophisticated means of encoding primitives and compounds which is capable of encoding with it internal structure. For example, for a given word W, we may speak of a double (S, T), where S is an finite sequence of natural numbers (representing semantic primes), and T is a binary tree. We might encode S using the Chinese Remainder Theorem [37], and T using any bijective mapping of the set of binary trees over $\mathbb{N}$. The numbers which encode S and T could then in turn be encoded into a single number, W, using a pairing function[28]. While the elegance of the system of prime number coding described by Rhiemeier stems from its simplicitly, it appears to necessarily come at the cost of semantic imprecision.

[28]Unfortunately, there is some redundancy in this scheme. Ideally, for any number n such that $n \in \mathbb{N}$, there is one and only one decomposition into (S, T) and vice versa. While the former is true for the described encoding, the latter is not: there are numbers with no valid decoding. This is because the length is effectively encoded twice. We encode the number of primes explicitly in our coding of S, and implicitly in mapping T to $\mathbb{N}$. The redundancy occurs because for any given tree with a given amount of leaf nodes (i.e. nodes without child nodes), only sequences with a matching length can be validly paired.

2.12 Reconstructions

Proto-Indo-European

References

[1] Torben Andersen. "The Phonemic System of Agar Dinka". In: *Journal of African Languages and Linguistics* 9.1 (1987). Ed. by Felix K. Ameka and Azeb Amha, pp. 1–28 (cit. on p. 19).

[2] Gregory D. S. Anderson. "Burushaski Morphology". In: *Morphologies of Asia and Africa*. Ed. by Alan S. Kaye. Vol. 2. Eisenbrauns, 2007. Chap. 44, pp. 1233–1275 (cit. on p. 25).

[3] Juliette Blevins and Sheldon P. Harrison. "Trimoraic Feet in Gilbertese". In: *Oceanic Linguistics* 38.2 (Dec. 1999), pp. 203–230. ISSN: 00298115, 15279421. URL: http://www.jstor.org/stable/3623292 (cit. on p. 27).

[4] Nancy Bowers. "Kapauku Numeration: Reckoning, Racism, Scholarship, and Melanesian Counting Systems". In: *The Journal of the Polynesian Society* 86.1 (1977), pp. 105–116. ISSN: 00324000. URL: http://www.jstor.org/stable/20705230 (cit. on p. 27).

[5] Humphrey Carpenter. *J.R.R. Tolkien: A Biography*. 1st ed. New York, NY, USA: Houghton Mifflin, 2000 (cit. on p. 35).

[6] Sonja Fritz. *The Dhivehi Language. A Descriptive and Historical Grammar of Maldivian and Its Dialects*. Trans. from the German by Sonja Fritz. Vol. 1. 2 vols. Beiträge zur Südasienforschung 191. Heidelberg: Ergon-Verlag, 2002. 270 pp. ISBN: 9783899132489 (cit. on pp. 21–23).

[7] Julia Dieterman van Haitsma and Willard van Haitsma. *A Hierarchical Sketch of Mixe. As Spoken in San José El Paraíso*. Ed. by Irvine Davis. SIL Publications in Liguistics and Related Fields 44. Summer Institute of Linguistics, Feb. 1976 (cit. on pp. 29–32).

[8] Richard K. Harrison. *Sonagona*. Sona Language Study Circle. 2013. URL: http://www.sonagona.org/index.html (visited on 09/22/2016) (cit. on p. 41).

[9] Sheldon P. Harrison. "Linguistic Evidence for Polynesian Influence in the Gilbert Islands". In: *Language Contact and Change in the Austronesian World*. Trends in Linguistics. Ed. by Thomas Edward Dutton and Darrell T. Tryon. Studies and Monographs 77. Berlin: Walter de Gruyter, 1994, pp. 321–349. ISBN: 9783110127867 (cit. on p. 27).

[10] Martin Haspelmath. *Explaining Alienability Contrasts in Adnominal Possession: Economy vs. Iconicity*. Lecture Handout. Second Conference on the Syntax of the World's Languages: University of Lancaster, Sept. 14–17, 2006. URL: http://email.eva.mpg.de/~haspelmt/2006swl.pdf (cit. on p. 21).

[11] Elizabeth Hume-O'Haire and Mike Armstrong. *Fur*. Metathesis in Language Database. Ohio State University. June 13, 2007. URL: http://metathesisinlanguage.osu.edu/LanguageDetails.cfm?LanguageID=48 (cit. on p. 19).

[12] Ronald I. Kim. *Introduction to Tocharian*. Lecture Notes. Institute for Comparative Linguistics: Charles University of Prague, Nov. 19–23, 2012. URL: http://enlil.ff.cuni.cz/system/files/tocharian.pdf (cit. on p. 25).

[13] Todd B. Krause and Jonathan Slocum. *Tocharian Online*. Ed. by Jonathan Slocum and Winfred P. Lehmann. University of Texas. May 2014. URL: http://www.utexas.edu/cola/centers/lrc/eieol/tokol-0.html (cit. on pp. 23, 25).

[14] Johanna Nichols and Balthasar Bickel. "Possessive Classification". In: *The World Atlas of Language Structures Online*. Ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. URL: http://wals.info/chapter/59 (cit. on p. 26).

[15] Erin O'Rourke. "Dialect Differences and the Bilingual Vowel Space in Peruvian Spanish". In: *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*. 4th Conference on Laboratory Approaches to Spanish Phonology. (University of Texas at Austin, Sept. 26–28, 2008). Ed. by Marta Ortega-Llebaria. Somerville, MA, USA: Cascadilla Proceedings Project, 2010, pp. 20–30 (cit. on p. 42).

[16] David J. Peterson. *About David J. Peterson*. Dothraki Official Website. 2016. URL: http://www.dothraki.com/about-david-j-peterson/ (visited on 09/20/2016) (cit. on p. 48).

[17] David J. Peterson. *Dothraki: A Language of Fire and Blood*. Dothraki Official Website. 2016. URL: http://www.dothraki.com/ (visited on 09/20/2016) (cit. on p. 48).

[18] David J. Peterson. *Perzo Vūjita*. Dothraki Official Website. 2016. URL: http://www.dothraki.com/2013/04/perzo-vujita/ (visited on 09/20/2016) (cit. on p. 48).

[19] David J. Peterson. *Sesīr Urnēbion Zȳhon Keliton Issa*. Dothraki Official Website. 2016. URL: http://www.dothraki.com/2013/04/sesir-urnebion-z%C8%B3hon-keliton-issa/ (visited on 09/21/2016) (cit. on p. 50).

[20] David J. Peterson. *Tȳni Trēsi*. Dothraki Official Website. 2016. URL: http://www.dothraki.com/2013/05/t%C8%B3ni-tresi/ (visited on 09/20/2016) (cit. on p. 50).

[21] David J. Peterson. *Valyrian Adjectives*. Dothraki Official Website. 2016. URL: http://www.dothraki.com/2013/07/valyrian-adjectives/ (visited on 09/21/2016) (cit. on p. 50).

[22] Kenneth L. Pike. "Language and Life. Part III: A Training Device for Translation Theory and Practice". In: *Bibliotheca Sacra* 114.456 (Oct.–Dec. 1957), pp. 347–362 (cit. on pp. 35–37).

[23] Jeff Prothero. *Design and Implementation of a Near-optimal Loglan Syntax*. May 9, 1990. URL: https://mw.lojban.org/papri/Design_and_Implementation_of_a_Near-optimal_Loglan_Syntax (visited on 09/19/2016) (cit. on p. 51).

[24] Bert Remijsen. "Tonal alignment is contrastive in falling contours in Dinka". In: *Language* 89.2 (2013), pp. 297–327 (cit. on p. 19).

[25] Jörg Rhiemeier. *Arithmographic Language*. FrathWiki. Nov. 17, 2011. URL: http://www.frathwiki.com/Arithmographic_language (visited on 09/17/2016) (cit. on p. 53).

[26] Jörg Rhiemeier. *Designing an International Auxiliary Language*. Jörg Rhiemeier's Conlang Page. Nov. 17, 2007. URL: http://www.joerg-rhiemeier.de/Conlang/auxlang-design.html (visited on 09/27/2016) (cit. on p. 42).

[27] Jörg Rhiemeier. *Old Albic*. FrathWiki. June 17, 2015. URL: http://www.frathwiki.com/Old_Albic (visited on 09/08/2016) (cit. on pp. 38, 39).

[28] Jörg Rhiemeier. *X-1*. FrathWiki. Oct. 31, 2014. URL: http://www.frathwiki.com/X-1 (visited on 09/17/2016) (cit. on pp. 51, 52).

[29] Jörg Rhiemeier. *X-5*. FrathWiki. July 17, 2011. URL: http://www.frathwiki.com/X-5 (visited on 09/17/2016) (cit. on p. 53).

[30] Pierric Sans. *Does Bésiro have a 'quadripartite' alignment?* SSILA Winter Meeting. Portland, OR, Jan. 5–8, 2012. URL: https://www.academia.edu/2178570/ (cit. on p. 26).

[31] Yair Sapir. "Elfdalian. The Vernacular of Övdaln". In: *Proceedings of the First Conference on Övdalian*. The First Conference on Övdalian. (June 18–19, 2004). Uppsala University. Älvdalen, Sweden (cit. on pp. 20, 21).

[32] Kenneth Searight. *Sona: An Auxiliary Neutral Language*. With an intro. by Charles Kay Ogden. Psyche Miniatures 75. London: K. Paul, Trench, Trubner & Company, 1935. 119 pp. (cit. on pp. 41–43).

[33] Gary F. Simons, ed. *Ethnologue*. Version 19. SIL International. May 2016. URL: http://www.ethnologue.com (cit. on pp. 7, 26, 27).

[34] Andrew Spencer. "Bracketing Paradoxes and the English Lexicon". In: *Language* 64.4 (1988), pp. 663–682 (cit. on p. 53).

[35] Ishaan Tharoor. *Tongues of Ice and Fire: Creating the Languages in* Game of Thrones. Time Magazine. May 3, 2013. URL: http://entertainment.time.com/2013/05/03/tongues-of-ice-and-fire-creating-the-languages-of-game-of-thrones/ (visited on 09/20/2016) (cit. on p. 48).

[36] Zachary Weaver. *Set Semantics in Davin*. Lecture Slides. 5th Language Creation Conference. Austin, Texas: University of Texas, May 4–5, 2013. URL: http://conlang.org/cms/wp-content/uploads/zweaver_lcc5_slides.pdf (cit. on pp. 34, 35).

[37] Eric W. Weisstein. *Chinese Remainder Theorem*. Wolfram MathWorld. Aug. 18, 2016. URL: http://mathworld.wolfram.com/ChineseRemainderTheorem.html (visited on 09/18/2016) (cit. on p. 53).

[38] Eric W. Weisstein. *Fundamental Theorem of Arithmetic*. Wolfram MathWorld. Aug. 18, 2016. URL: http://mathworld.wolfram.com/FundamentalTheoremofArithmetic.html (visited on 09/18/2016) (cit. on p. 53).

[39] John Wilkins. *An Essay Towards a Real Character and a Philosophical Language*. London: Gellibrand, 1668 (cit. on p. 53).

Part II

Grammar

Chapter 3

Phonology

Aa	Ææ	Bb	Cc	Dd	Ðð	Ee	Ëë
Ff	Gg	Gg̱	Hh	Ĥĥ	Hʉʊ	Ii	ȝȝ
Jj	Kk	Ll	Ḷḷ	Mm	Ṃṃ	Nn	Ṇṇ
Oo	Øø	Ǫǫ	Œœ	Pp	Rr	Ṛṛ	Ss
Tt	Uu	Üü	Vv	Ww	Yy	Þþ	Ƿƿ

Figure 3.1: The Dunsish Alphabet

	Labial		**Alveolar**		**Palatal**		**Velar**		**Glottal**	
Nasal	ṃ	m	ṇ	n						
Stop	p	b	t	d			k	g		
Sibilant				s						
Fricative	f	v	þ	ð	c				h	ĥ
Approximant						j	ʰʊ	ƿ		
Lateral			ḷ	l						
Trill			ṛ	r						

Figure 3.2: Consonant Phonemes

3.1 Vowels

Vowels have three pertinent qualities: length, pitch, nasalization.

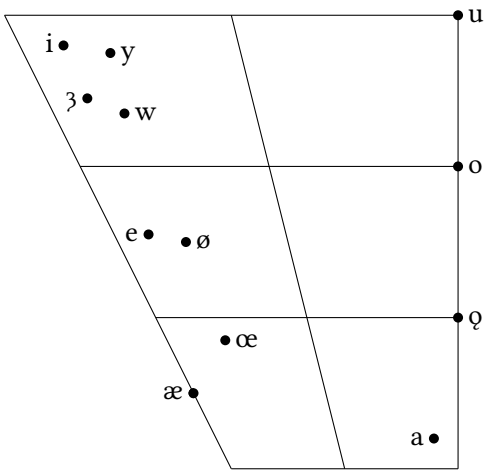


Figure 3.3: Vowel System

№	Front	Round	Back	Diphthong
1	i	y	u	au
2	ɜ	w	o	ɒu
3	e	ø	ɒ	ei
4	æ	œ	a	æi

Table 3.1: Ablaut System

Length

Like Dinka (and possibly Estonian), Dunsish distinguishes between three different lengths of vowels, as opposed to the usual two.

Pitch

Dunsish features a pitch accent system similar to that of Japanese or Ancient Greek. Pitch can occur on any vocalic mora.

Nasalization

Vowels which occur before /n/ have their last mora nasalized. All subsequent vowels feature vowel harmony with this, and thus are nasalized. Vowel harmony does not cross stem-boundaries.

3.2 Allophony

Phonological Processes

Following the generative paradigm, the phonological processes of Dunsish can be described in a few simple rules, using an EBNF (Extended Backus-Naur Form) like metasyntax:

1. $V \rightarrow V_H\ V_L$ *Vowels*
2. $V_S \rightarrow V$ *Short Vowels*
3. $V_M \rightarrow V\text{:}$ *Medium Vowels*
4. $V_X \rightarrow V\text{::}$ *Long Vowels*
5. $C \rightarrow C_P\ C_F\ C_S$ *Consonants*
6. $C_P \rightarrow C_{Pv}\ C_{Pu}$ *Plosives (Stops)*
7. $C_{Pu} \rightarrow$ /p t k/ *Unvoiced Stops*
8. $C_{Pv} \rightarrow$ /b d g/ *Voiced Stops*
9. $C_F \rightarrow C_{Fv}\ C_{Fu}$ *Fricatives*
10. $C_{Fu} \rightarrow$ /f þ s x h/ *Unvoiced Fricatives*
11. $C_{Fv} \rightarrow$ /v ð q h/ *Voiced Fricatives*
12. $C_S \rightarrow C_{Sv}\ C_{Su}$ *Sonorants*
13. $C_{Su} \rightarrow$ /ṁ ṅ ṙ ḷ/ *Unvoiced Sonorants*
14. $C_{Sv} \rightarrow$ /m n r l/ *Voiced Sonorants*
15. $V_H \rightarrow$ /i u ü y w ẅ/ *High Vowels*
16. $V_L \rightarrow$ /e o ø æ a ä/ *Low Vowels*
17. $kV_H \rightarrow [c]V_H$ *Stop Palatization*
18. $n[c] \rightarrow [ɲc]$ *Nasal Palatization*
19. $(V \mid V_M)C_{Pu} \rightarrow V^hC_{Pu}$ *Pre-aspiration*

Breathy Voice

Intervocalically, the voiceless liquids (/ṁ/ /ṅ/ /ḷ/ /ṙ/) become "breathy" voiced, or murmured. Thus, kwŵḷon is pronounced /koːlʱn̥n/.

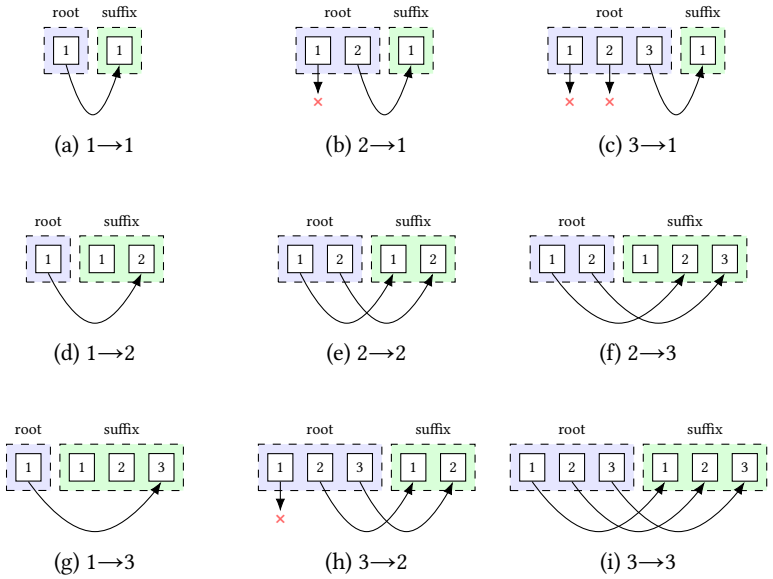


Figure 3.4: Tone Stealing

Aspiration

Following short vowels, voiceless stops become pre-aspirated as in Faroese or Icelandic[1]. It should be noted that stops are *not* normally (post) aspirated at the beginning of words as they are in English.

Tone Stealing

Dunsish undergoes a phenomenon called "tone stealing." In essence, tone stealing is a phenomenon by which pitch accent is "stolen" from the root vowel and transfered to the suffix vowel, as per figure 3.4.

[1]Note that Árnason [1: p. 219] presents evidence to suggest that pre-aspiration in Icelandic is not "true" pre-aspiration, but the insertion of the /h/ phone before stops. Nevertheless, it appears that Faroese does in fact have true pre-aspiration.

References

[1] Kristján Árnason. *The Phonology of Icelandic and Faroese*. Oxford Linguistics. OUP Oxford, 2011. ISBN: 9780199229314 (cit. on p. 64).

Chapter 4

Number

4.1 Introduction

Dunsish features a complex number system.

4.2 Nullary

The nullary number denotes zero amount of an item.

4.3 Singular

The singular represents a single object.

4.4 Ambal

The ambal is used to denote a natural pairing or amount of an object. Like Tocharian, Dunsish distinguishes between a dual, representing any two objects, and an ambal for denoting a natural grouping [3]. Unlike Tocharian, in Dunsish, the ambal denotes a natural grouping of any amount, not just pairs of two[1].

4.5 Paucal

The paucal is used to refer to "a few" or small number of objects.

4.6 Plural

The plural is used to refer to more than three objects.

[1] That Tocharian distinguished between a dual and a paral for casual and natural pairs respectively is disputed by Kim [2: p. 26].

4.7 Dual

The dual denotes two of an object.

4.8 Triple

The triple number refers to three of an object.

4.9 Indefinite

The indefinite refers to an unspecific number of objects, i.e. 1 or more. The semantics of the indefinite number are identical to that of Basque [1: p. 31]. The indefinite is the unmarked/default form.

4.10 Fractional

The fractional denotes a partial amount of an object.

4.11 Numeral

The numeral number is used when a specific amount is specified

References

[1] Rudolf P.G. de Rijk. *Standard Basque. A Progressive Grammar*. Current Studies in Linguistics. MIT Press, Nov. 2007 (cit. on p. 68).

[2] Ronald I. Kim. *Introduction to Tocharian*. Lecture Notes. Institute for Comparative Linguistics: Charles University of Prague, Nov. 19–23, 2012. URL: http://enlil.ff.cuni.cz/system/files/tocharian.pdf (cit. on p. 67).

[3] Todd B. Krause and Jonathan Slocum. *Tocharian Online*. Ed. by Jonathan Slocum and Winfred P. Lehmann. University of Texas. May 2014. URL: http://www.utexas.edu/cola/centers/lrc/eieol/tokol-0.html (cit. on p. 67).

Chapter 5

Lexical Semantics

Underlying Dunsish is a systematic lexical semantic system, which affords for a precise means of accurately deriving new words in a predictable manner and makes the relationships between lexemes easy to understand.

5.1 René de Saussure

One of the earliest works that might be described as concerned with lexical semantics was the attempts of René de Saussure to establish a rigorous account of the semantics of Esperanto. Brother of the more famous Ferdinand de Saussure, René de Saussure was a mathematician and early exponent of Esperanto. He developed a unique theoretical machinery wherein each Esperanto root was said to have an inherent character: object, process, or quality [1: pp. 150–152].

References

[1] P.G. Forster. *The Esperanto Movement*. Contributions to the sociology of language. Mouton, 1982. ISBN: 9789027933997 (cit. on p. 71).

Chapter 6

Gender System

Dunsish features two different gender systems: *element* and *class*.

6.1 Element

Each Dunsish lexeme belongs to one of five *elements*: earth, water, air, fire, or ether. Corbett distinguishes between *formal* and *semantic* systems of gender assignment [1]. Semantic systems of assignment are those which assign gender based on qualities such as sex or animacy. Conversely, formal gender assigns words arbitrarily. General speaking, it is rare (or in the case of formal gender, non-existent) to find a gender system which is "pure." Nevertheless, we can distinguish between a few different categories.

Semantic

A famous example of semantic gender[1] is the word class system of the Bantu languages [2]. There are around 23 different classes found among the Bantu languages, though no single languge has each of them[2].

Quasi-Semantic

Quasi-semantic is what I term the gender system often found in Indo-European languages.

Idiosemantic

Idiosemantic is a word of my own coining which I use to refer to the gender system found in Dunsish (and possibly other languages). In essence, an idiosemantic system is one in which gender assigment is assigned according to the semantics of the word, but

[1] Sometimes, the term "word class" is used. However, the boundary between *word class* and *gender* is not always clear.

[2] Though Proto-Bantu itself is reconstructed as having 23.

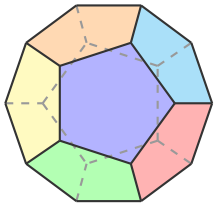
the categorization is based on a symbolic, metaphorical, or otherwise idiosyncratic criteria. So while in theory the gender assignment is predictable based off of the meaning of a particular word, in practice, it must be memorized.

(a) Earth

(b) Water

(c) Air

(d) Fire



(e) Ether

Figure 6.1: The Five Elements

References

[1] Greville G. Corbett. "Systems of Gender Assignment". In: *The World Atlas of Language Structures Online*. Ed. by Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. URL: `http://wals.info/chapter/32` (cit. on p. 73).

[2] Katherine Demuth. "Bantu Noun Class Systems: Loanword and Acquisition". In: *Systems of Nominal Classification*. Ed. by Gunter Senft. Vol. 4. Language Culture and Cognition. Cambridge University Press, 2008, pp. 270–292 (cit. on p. 73).

Chapter 7

Case Morphology

Dunsish features a complex system of case marking, almost totally eschewing the use of adpositions, found in more isolating languages such as English, or even the Romance languages. The case system is inspired by that of Tocharian, which features a "tiered," or semi-agglutinative, system of marking [1]. Dunsish cases are divided into several categories: *core*, *oblique*, *nominal*, and *spatio-temporal*.

7.1 Morphosyntactic Alignment

Before enumerating a list of cases, a discussion of morphosyntactic alignment and argument structure is in order

Thematic Roles

Agent

An agent is a willed initiator of an action upon some object or individual.

Co-agent

A co-agent is a participant in a willed, reciprocal action. It is important to understand that a co-agentive role is one in which the action performed is *intrinsically* reciprocal. To illustrate this difference, considering the following sentences:

(3) Alice and Bob hit each other.

(4) Alice hit Bob, and Bob hit Alice.

(5) Alice married Bob.

(6) * Alice married Bob, and Bob married Alice.

As can be seen, sentences 3 and 4 are essentially paraphrases of one another. Or, in other words, me might say that sentence 3 is merely an abbreviated form of 4. In any case, the point to be made here is that the semantics of *hitting* do not contain any inherent notion of reciprocality. Converseley, sentence 6, while not unclear in

meaning, is either ungrammatical, or at least needlessly redundant in form. We can see that to marry someone is an inherently reciprocal action, which requires two participants.

Force

A force is an inanimate cause of some process which acts upon the patient.

Patient

A patient (animate or inanimate) is that which a process or action is acted upon.

Theme

A theme is that which modulates another process.

Recipient

A recipient is that which receives the theme from the agent.

Ditransitive Constructions

Objects

Ditransitives are deceptively complex. Consider the following English examples:

(7) Alice gave Bob the ball. *Theme-like*

(8) Alice threw Bob the ball. *Patient-like*

Sentences 7 and 8 are, in English, syntactically identical, but the semantic role each of the three participants (Alice, Bob, the ball) plays in these sentences is quite different. To merely *give* an object to someone does not, in itself, imply that any physical action is imparted upon the gift itself. In fact, it is perfectly intelligible to speak of giving (in the sense of transferring ownership or possesion of) objects which are not physically present. While to speak of giving might pragmatically *suggest* that one is physically handing over some material thing, it does not necessitate that this is the case. Conversely, the implications of throwing an object are orthogonal to that of giving.

(9) * Alice threw Bob the house.

Thus, as we can see, sentence 9 is nonsensical. In sentence 8, what is most salient is that Alice performed an action (i.e. throwing) upon the ball. The recipient, Bob, modulates Alice's activity: it provides further detail on what Alice did to the ball. Furthermore, that Alice threw to ball *to* Bob does not in anyway imply that she *gave* the ball to Bob (though obviously it suggests that Bob is now in physical possesion of the ball).

Recipients

Let us now consider the difference roles in which syntactic recipients may function.

(10) Alice gave Bob the flu. *Patient-like*

(11) Alice gave Bob the money. *Recipient-like*

As with previous examples, the syntactic similarity between sentences 10 and 11 obscures the important semantic differences between the two. To illustrate this difference, consider that sentence 10 is roughly semantically equivalent to sentence 12:

(12) Alice made Bob sick (with the flu).

Unlike the previous examples, sentence 12 has no (syntactic) indirect object. Alice functions as the agent, and Bob as the patient[1]. In other words, the (semantic) argument structure of sentence 10 is opposite to that of 8. In sentence 10, it is the theme, the flu, which modulates Alice's activity upon the syntactic recipient, but semantic patient, Bob.

7.2 Core Cases

Dunsish has a so-called "quadripartite" morphosyntactic alignment, based off of research on the Chiquitano language of Bolivia [3]. Figure 7.1 shows representations of five possible morphosyntactic alignments[2]. In addition, Dunsish features a so-called secundative alignment[3], which refers to the fact that the recipient of a ditransitive verbs is treated as the patient, and the transfered object is treated as the theme [2].

Ergative

The ergative case (glossed ERG) is used for the subject of transitive verbs.

Accusative

The accusative case (glossed ACC) is used for the object of transitive verbs.

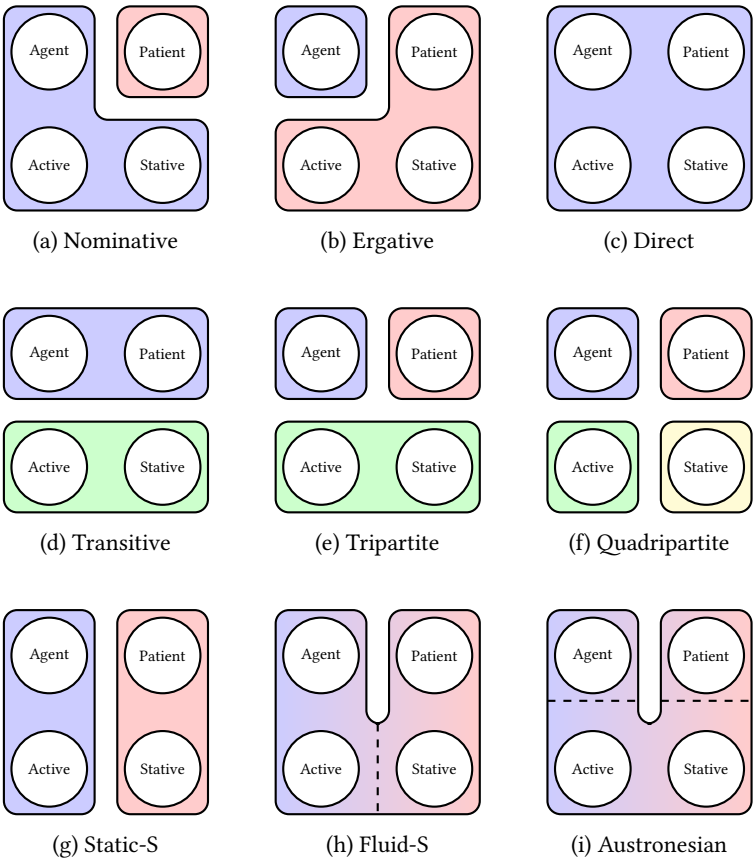
Nominative

The nominative case (glossed NOM) is used for the subject of active-type intransitive verbs.

[1]It is worth noting that, since English does not mark for volition, there are two possible readings of sentence 12. We might interpret it to mean either that Alice intentionally caused Bob to become sick, or that her presence led to Bob contracting the disease. For the sake of this discussion, we will consider only the former.

[2]Note that tripartite alignment (7.1e) is sometimes known as ergative-accusative, and active-stative alignment (7.1g) as split intransive.

[3]Also known as "dechticaetiative" or primary object language



(a) Nominative (b) Ergative (c) Direct

(d) Transitive (e) Tripartite (f) Quadripartite

(g) Static-S (h) Fluid-S (i) Austronesian

Figure 7.1: Morphosyntactic Alignments

Absolutive

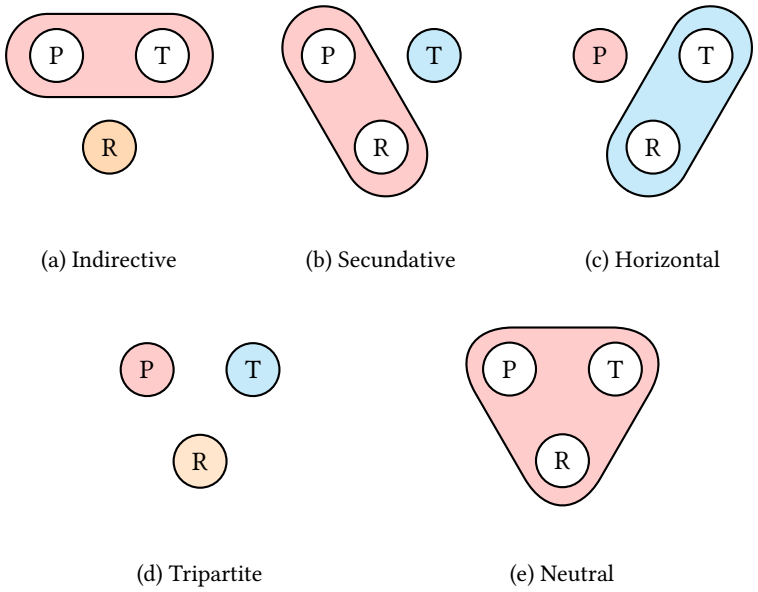
The absolutive case (glossed ABS) is used for the subject of stative-type intransitive verbs.

Verbal

In addition to the four core "argument" cases, there is a case for what in other languages would hold the verb: the action of the statement. This works somewhat similar to relationals in Kēlen [4].

7.3 Polytransitive Cases

Next, Dunsish has a set of cases which I call "polytransitive." The polytransitive cases are who which function as part of the valency increasing operations and/or ditransitive constructions. Figure 7.2 shows illustrations of the possible ditransitive alignments, in analogy to the more general idea of morphosyntactic alignment[4].



(a) Indirective (b) Secundative (c) Horizontal

(d) Tripartite (e) Neutral

Figure 7.2: Ditransitive Alignments

[4]The terms "neutral," "horizontal," and "tripartite" (in this context) are taken from Mal'chukov et al. [2].

References

[1] Todd B. Krause and Jonathan Slocum. *Tocharian Online*. Ed. by Jonathan Slocum and Winfred P. Lehmann. University of Texas. May 2014. URL: http://www.utexas.edu/cola/centers/lrc/eieol/tokol-0.html (cit. on p. 77).

[2] Andreĭ L'vovich Mal'chukov, Martin Haspelmath, and Bernard Comrie. "Ditransitive constructions: A Typological Overview". In: *Studies in Ditransitive Constructions: A Comparative Handbook*. De Gruyter Mouton, 2010. Chap. 1, pp. 1–64. ISBN: 9783110220360 (cit. on pp. 79, 81).

[3] Pierric Sans. *Does Bésiro have a 'quadripartite' alignment?* SSILA Winter Meeting. Portland, OR, Jan. 5–8, 2012. URL: https://www.academia.edu/2178570/ (cit. on p. 79).

[4] Sylvia Sotomayor. *An Introduction to Kēlen*. Aug. 2011. URL: http://www.terjemar.net/kelen.php (cit. on p. 80).

Chapter 8

Tense, Aspect, Mood

8.1 Introduction

Unlike Indo-European language, Dunsish marks only for aspect and mood, while "tense" marking is acheived by means of a set of temporal cases. Dunsish's tense, aspect, and mood (TAM) system is inspired by Reichenbach's theory of tense, Prior's temporal logic, and Allen's set of temporal relations.

8.2 Allen Relations

In [1], Allen introduced a calculus for describing the relationships between two (or possibly more) intervals of time. The set of 13 relations are exhaustive: they describe every possible relation.

References

[1] James F. Allen. “Maintaining Knowledge About Temporal Intervals”. In: *Commun. ACM* 26.11 (Nov. 1983), pp. 832–843. ISSN: 0001-0782. DOI: 10.1145/182.358434. URL: http://doi.acm.org/10.1145/182.358434 (cit. on p. 83).

Part III

Appendix

Appendix A

Old Norse Corpus Linguistics

A.1 Introduction

Inkeeping with my aesthetic tastes, I have endeavored to design Dunsish to, in large part, emulate the sound of Old Norse[1]. Unfortunately, I was unable to find any data on those phonological qualities of the Old Norse language that were relevant my particular interests. As such, I have performed a brief study of an Old Norse corpus in order to extract and quantify the salient qualities of Old Norse which are salient to phonological design of Dunsish.

Unfortunately, what we know of Old Norse by necessity comes to us only in the form of the rarefied literary register employed by the likes of the Snorri Sturluson and the authors of the Sagas. However, as I am mostly interested in reproducing a particular sound that is pleasing to my ear, and not (necessarily) the spoken vernacular, this is of secondary concern. Nevertheless, it is worth pointing out that students of Old Norse phonology are considerably luckier than many of their fellow philologist, as we have available to our use works such as the famous *First Grammatical Treatise* which treats the subject of then contemporary speech with laudable rigor that is in many ways prescient of modern linguistic methodology.

A.2 Corpus

Since we do not have available to use living Old Norse speakers, in order to study Old Norse phonology we must, as mentioned, resort to surviving literary texts. As such, I have assembled a corpus of Old Norse text from works hosted on the Perseus Project[2]. The Perseus Project provides an XML file of each text to download. Using a Perl script, I postprocessed these these XML files by stripping away markup, punctuation, etc. and then concatenated the contents of each file together to produce one single file consisting of the bare words from each document. The final corpus contains 421,056 words.

[1] Though not entirely. Dunsish's pitch accent system (whether such a category is tenable) is inspired by that of Greek and Japanese. Furthermore, it's vowel length system is a direct borrowing from Dinka.

[2] Available at `http://www.perseus.tufts.edu/hopper/`.

A.3 Statistics

Sound Frequency

After generating this corpus, my next goal was to run various statistical tests. Since Old Norse orthography is largely phonemic, a naïve test of the relative frequency of letter-forms proved far more useful than it would have in studying languages such as modern English or French. The phonological character is of course a function not just of the phonological inventory of a particular language, but also, crucially, of the relative frequency of those sounds which compose this inventory. The results of this test are given in table A.1.

Letter	Sound	% Cons.	% Vowel	Freq.
r				
a				
n				
i				
e				
t				
s				
k				
l				
u				
m				
g				
ð				
h				
o				
v				
þ				
f				
á				
d				
í				
ó				
j				
b				
ǫ				
y				
æ				
p				
ú				
é				
ý				
z				
x				
c				
q				

Table A.1: Frequency of Phones

Appendix B

Lexicon Format

This appendix contains documentation for the custom XML-based format which I have devised to manage and catalogue the lexicon of my conlang Dunsish.

B.1 Introduction

Initially, I had planned to use an SQL database to manage Dunsish's lexicon. However, this option was quickly scrapped for a number of reasons. First, interoptability with LuaTEX is of utmost concern, and based on my research, setting up the Lua SQL libraries is tedious and error-prone. Second, I would like all my documents to require as few external tools to compile as possible, and thus maintain maximal portability. Finally, integrating SQL with version control, while possible, can be rather obnoxious.

In light of these issues, I elected instead to design a custom format based on XML. While XML has been criticized for poor readability, the situation is certainly superior to that of SQL. Though XML is undoubtedly abused, document-like objects such as lexicons or dictionaries are where its use is most appropriate. In fact this particular application has a long pedigree: the Oxford English Dictionary was the among the earliest adopters of SGML, the predecessor of XML [1].

B.2 Methodology

Initially, I attempted to specify the lexicon format using the W3C's XML Schema language (XSD) [4]. After using XSD for a bit, I quickly decided to look for other options. Strangely, despite being an official W3C recommendation, XSD 1.1 has exceedingly poor adoption. The `xmllint` utility from the libxml2 project does not support it, and as far as I know, there are no open source tools which do. In any case, XSD is an exceedingly ugly and tedious language to deal with. Instead, I chose to go with the much more pleasant RELAX NG schema language [3]. While it lacks the same expansive feature set as XSD 1.1, this is essentially a moot point given the aforementioned. As a surrogate, I elected to use the Schematron language in order to perform further data validation needs beyond the capabilities of RELAX NG, or for that matter, XSD 1.0 [2].

While the circumstances which led to me adopting this combination were rather accidental, it was in a way fortuitous. The RELAX NG schema specifies the structure of the lexicon document, while the schematron enforces so-called co-constraints, and ensures that the document is semantically sound. For other conlangers similarly looking for a way to catalogue their lexicon (or other linguistic data), I highly recommend this toolset. The alternative "compact" syntax that is available for RELAX NG is, if nothing else, a marked improvement over the laughably bloated XSD.

An obvious question might be whether bothering with a schema is even worthwhile. After all, if we're talking about conlangs, it's unlikely that the lexicon will be worked on by any more than even a handful of people. The answer is that even for conlangs developed solely by a single creator, ensuring the lexicon is written in a consistent manner can require much mental overhead. The advantage of a scheme is absolute assurance that your document adheres to a coherent model.

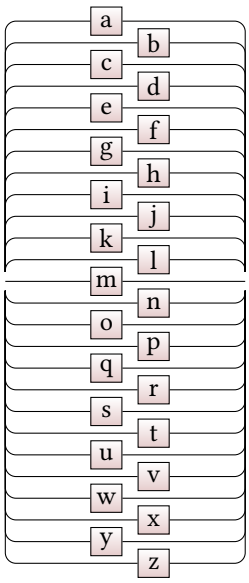
B.3 Design

The Dunsish lexicon format, like all XML documents, contains a single root node: `lexicon`, which has no attributes. Inside the `lexicon` node there are four types of lexical entries: `root`, `antiroot`, `infix`, and `stem`.

Entries

Common to all types of entries are the `id` attribute and `def` and `note` elements. The attribute `id` contains a unique indentifier for entry cross references in `antiroot` and `stem` entries. The `note` element contains any supplemental documentation, not intended to be included in the final LaTeX output[1]. The regular expression which denotes the set of valid values for the `id` attribute is shown in Figure B.1.

[1] I do intend, however, to make this information available in the web version produced by an XSLT stylesheet, but this is only for my own uses.



`[a-z](_?[a-z0-9])*`

Figure B.1: `id` syntax diagram

References

[1] Laura Elliot. "How the Oxford English Dictionary Went Online". In: *Ariadne* 24 (June 21, 2000). URL: http://www.ariadne.ac.uk/issue24/oed-tech/ (visited on 02/02/2017) (cit. on p. 91).

[2] Rick Jelliffe. *Schematron*. 2017. URL: http://schematron.com/ (visited on 02/02/2017) (cit. on p. 91).

[3] Murata Makoto. *RELAX NG*. Feb. 25, 2014. URL: http://relaxng.org/ (visited on 02/02/2017) (cit. on p. 91).

[4] C. M. Sperberg-McQueen and Henry Thompson. *W3C XML Schema*. World Wide Web Consortium (W3C). Apr. 2000. URL: https://www.w3.org/XML/Schema (visited on 02/02/2017) (cit. on p. 91).

Appendix C

Music Theory

This chapter will contain information on the accompanying music theory of Dunsish.
My sources include [2] and [1].

References

[1] Dave Benson. *Music: A Mathematical Offering*. Web. New York: Cambridge University Press, Dec. 14, 2008. URL: https://homepages.abdn.ac.uk/mth192/pages/html/maths-music.html (cit. on p. 95).

[2] J. N. Hooker. "Finding Alternative Musical Scales". In: *Principles and Practice of Constraint Programming. Proceedings*. The 22nd International Conference on Principles and Practice of Constraint Programming. (Toulouse, France, Sept. 5–9, 2016). Ed. by Michel Rueher. Lecture Notes in Computer Science 9892. Association for Constraint Programming. Switzerland: Springer International Publishing, 2016, pp. 753–768. ISBN: 978-3-319-44953-1. DOI: 10.1007/978-3-319-44953-1_47. URL: http://dx.doi.org/10.1007/978-3-319-44953-1_47 (cit. on p. 95).

Appendix D

Unit System

Dunsish has a corresponding unit system which is peculiar to it.

D.1 Length

The Dunsish unit of length is equivalent to the wavelength of the 21 cm line of Hydrogen-1.

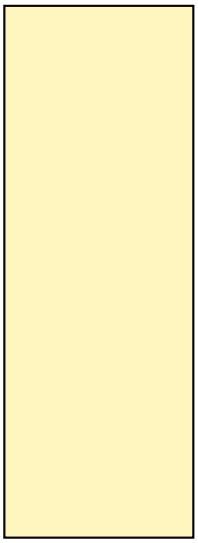


Figure D.1: Dunsish Unit of Length

Appendix E

QXZ Encoding

QXZ is an ASCII-based encoding for Dunsish text. The encoding is similar in spirit to the so-called x-system and h-system used for Esperanto [2], as well as Jim Henry's ASCII-based orthography for gjâ-zym-byn [1].

QXZ is named for the fact that it makes use of the Latin letters *Q*, *X*, and *Z*, which are not used in the full Dunsish alphabet. As such, there is thus no ambiguity in digraphs formed with these letters. The main benefit of QXZ is that it requires only the use of ASCII letters, and thus should be maximally usable in essentially any situation. I intentionally avoided the use of numerals and punctuation so as to allow for the use of QXZ-encoded strings in programming languages. Table E.1 shows the encoding in full. Note that the letters 'Q', 'X', and 'Z' as used in the encoding are case-insensitive, and a conformant implementation must treat e.g. `oXXQ` and `oxxq` identifically. However, so as to indicate case, all *other* characters are case-senstive.

Table E.1: QXZ Encoding

Letter	QXZ	Dunsish	Unicode	Postscript	TEX
A	A		U+0041	/A	
a	a		U+0061	/a	
Á	AQ		U+00C1	/Aacute	\'A
á	aq		U+00E1	/aacute	\'a
Æ	AX		U+00C6	/AE	\AE
æ	ax		U+00E6	/ae	\ae
Ǽ	AXQ		U+01FC	/AEacute	\'\AE
ǽ	axq		U+01FD	/aeacute	\'\ae
B	B		U+0042	/B	
b	b		U+0062	/b	
C	C		U+0043	/C	
c	c		U+0063	/c	
D	D		U+0044	/D	
d	d		U+0064	/d	

Continued on next page

Table E.1: QXZ Encoding — continued from previous page

Letter	QXZ	Dunsish	Unicode	Postscript	TEX
Ð	DX		U+00D0	/Eth	\DH
ð	dx		U+00F0	/eth	\dh
E	E		U+0045	/E	
e	e		U+0065	/e	
É	EQ		U+00C9	/Eacute	\'E
é	eq		U+00E9	/eacute	\'e
Ë	EX		U+00CB	/Edieresis	\"E
ë	ex		U+00EB	/edieresis	\"e
F	F		U+0046	/F	
f	f		U+0066	/f	
G	G		U+0047	/G	
g	g		U+0067	/g	
Ǥ	GX		U+01E4	/Gstroke	
ǥ	gx		U+01E5	/gstroke	
H	H		U+0048	/H	
h	h		U+0068	/h	
Ĥ	HX		U+0124	/Hcircumflex	\^H
ĥ	hx		U+0125	/hcircumflex	\^h
Ƕ	HZ	*hvæiṛ*	U+01F6		
ƕ	h		U+0195		
I	I		U+0049	/I	
i	i		U+0069	/i	
Í	IQ		U+00CD	/Iacute	\'I
í	iq		U+00ED	/iacute	\'i
ȝ	IX	*jȝga*	U+021C		
ȝ	ix		U+021D		
ȝ́	IXQ		U+021C U+0301		
ȝ́	ixq		U+021D U+0301		
J	J		U+004A	/J	
j	k		U+006A	/j	
K	K		U+004B	/K	
k	k		U+006B	/k	
L	L		U+004C	/L	
l	l		U+006C	/l	
Ḷ	LX		U+1E36	/Ldotbelow	\d{L}
ḷ	lx		U+1E37	/ldotbelow	\d{l}
M	M		U+004D	/M	
m	m		U+006D	/m	
Ṃ	MX		U+1E42	/Mdotbelow	\d{M}
ṃ	mx		U+1E43	/mdotbelow	\d{m}
N	N		U+004E	/N	
n	n		U+006E	/n	

Continued on next page

Table E.1: QXZ Encoding — continued from previous page

Letter	QXZ	Dunsish	Unicode	Postscript	TEX
Ṇ	NX		U+1E46	/Ndotbelow	\d{N}
ṇ	nx		U+1E47	/ndotbelow	\d{n}
O	O		U+004F	/O	
o	o		U+006F	/o	
Ó	OQ		U+00D3	/Oacute	\'O
ó	oq		U+00F3	/oacute	\'o
Ø	OX		U+00D8	/Oslash	\O
ø	ox		U+00F8	/oslash	\o
Ǿ	OXQ		U+01FE	/Oslashacute	\'\O
ǿ	oxq		U+01FF	/oslashacute	\'\o
Ǫ	OZ		U+01EA	/Oogonek	\k{O}
ǫ	oz		U+01EB	/oogonek	\k{o}
Ǫ́	OZQ		U+01EA U+0301		\k{\'O}
ǫ́	ozq		U+01EB U+0301		\k{\'O}
Œ	OXX		U+0152	/OE	\OE
œ	oxx		U+0153	/oe	\oe
Œ́	OXXQ		U+0152 U+0301		\'\OE
œ́	oxxq		U+0153 U+0301		\'\oe
P	P		U+0050	/P	
p	p		U+0070	/p	
R	R		U+0052	/R	
r	o		U+0072	/r	
Ṛ	RX		U+1E5A	/Rdotbelow	\d{R}
ṛ	rx		U+1E5B	/rdotbelow	\d{r}
S	S		U+0053	/S	
s	s		U+0073	/s	
T	T		U+0054	/T	
t	t		U+0074	/t	
U	U		U+0055	/U	
u	u		U+0075	/u	
Ú	UQ		U+00DA	/Uacute	\'U
ú	uq		U+00FA	/uacute	\'u
Ü	UX		U+00DC	/Udieresis	\"U
ü	ux		U+00FC	/udieresis	\"u
V	V		U+0056	/V	
v	v		U+0076	/v	
W	W		U+0057	/W	
w	w		U+0077	/w	
Ẃ	WQ		U+1E82	/Wacute	\'W
ẃ	wq		U+1E82	/wacute	\'w
Y	Y		U+0059	/Y	
y	y		U+0079	/y	

Continued on next page

Table E.1: QXZ Encoding — continued from previous page

Letter	QXZ	Dunsish	Unicode	Postscript	TEX
Ý	YQ		U+00DD	/Yacute	\'Y
ý	yq		U+00FD	/yacute	\'y
Þ	TX		U+00DE	/Thorn	\TH
þ	tx		U+00FE	/thorn	\th
Ƿ	WX		U+01F7	/Wynn	
ƿ	wx		U+01BF	/wynn	

References

[1] Jim Henry. *gjâ-zym-byn*. Dec. 2015. URL: http://jimhenry.conlang.org/gzb/gzb.htm (cit. on p. 99).

[2] Robert L. Read. *How to enter Esperanto text here*. Esperanto USA. Jan. 8, 2006. URL: https://www.esperanto-usa.org/en/node/108 (visited on 03/15/2017) (cit. on p. 99).

Appendix F

GNU Free Documentation License

Version 1.3
3 November 2008

Preamble

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

F.1 Applicability and Definition

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "**Document**", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "**you**". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "**Modified Version**" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "**Secondary Section**" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "**Invariant Sections**" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "**Cover Texts**" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "**Transparent**" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for auto-

matic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "**Opaque**".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "**Title Page**" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "**publisher**" means any person or entity that distributes copies of the Document to the public.

A section "**Entitled XYZ**" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "**Acknowledgements**", "**Dedications**", "**Endorsements**", or "**History**".) To "**Preserve the Title**" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

F.2 Verbatim Copying

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section F.3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

F.3 Copying in Quantity

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

F.4 Modification

You may copy and distribute a Modified Version of the Document under the conditions of sections F.2 and F.3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at

least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

F.5 Combining Document

You may combine the Document with other documents released under this License, under the terms defined in section F.4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

F.6 Collections of Document

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you

follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

F.7 Aggregation with Independent Works

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section F.3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

F.8 Translation

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section F.4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section F.4) to Preserve its Title (section F.1) will typically require changing the actual title.

F.9 Termination

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright

holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

F.10 Future Revisions of this License

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See `http://www.gnu.org/copyleft/`.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

F.11 Relicensing

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC,

and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.

Addendum: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

> Copyright © YEAR YOUR NAME. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with … Texts." line with this:

> with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

Appendix G

Unlicense

Preamble

While the document itself, and the text contained therein is available under the terms of the GNU Free Documentation License (GFDL, see Appendix F), the code used to generate it, and various helper scripts include in this repository I have placed under the terms of the Unlicense: essentially in the public domain.

Obviously, public domain material definitionally requires nothing on the part of those who make use of it. However, if you find any of my software useful, I would greatly appreciate some form of acknowledgement.

Unlicense Text

This is free and unencumbered software released into the public domain.

Anyone is free to copy, modify, publish, use, compile, sell, or distribute this software, either in source code form or as a compiled binary, for any purpose, commercial or non-commercial, and by any means.

In jurisdictions that recognize copyright laws, the author or authors of this software dedicate any and all copyright interest in the software to the public domain. We make this dedication for the benefit of the public at large and to the detriment of our heirs

and successors. We intend this dedication to be an overt act of relinquishment in perpetuity of all present and future rights to this software under copyright law.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

For more information, please refer to `http://unlicense.org`

Index

A Grammar of the Dunsish Language

by **Kevin Wesley Sīnā**

While there are hundreds, if not thousands, of conlangs (constructed language) scattered across the Internet, very few are documented in any appreciable quanity. With this book, I have attempted to document a fictional language, *Dunsish*, to the same degree of rigour and detail as might befit a natural language. In addition to a complete reference grammar of my conlang Dunsish, this book also contains:

- An extensive prolegomena containing overviews of:
  - 42 natural languages
  - 11 constructed languages
- A comprehensive bibliography
- A lexicon of Dunsish

It is my hope that this volume will prove instructive and, if nothing else, entertaining to those who read it.

This grammar was written in hope that it might be helpful and interesting to other conlangers. Thank you for reading. Support and feedback are appreciated!

About the Author

Kevin Sīnā is a software developer, based in New York City. His academic interests include Medieval logic, the Modist School, the Salamanca School and Baroque Scholasticism, Late Antiquity and Renaissance Platonism, utilitarian ethics, constructivist metaethics, and Pragmatism.

Bæcceitas Press

Suggested Donation

US	$5.00
AUS/CAN/NZ	$6.50
UK	£4.00
EU	€4.50