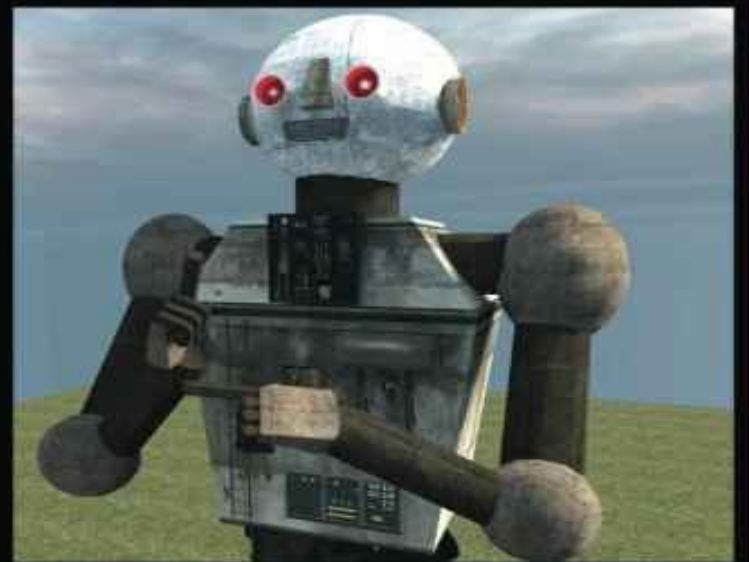
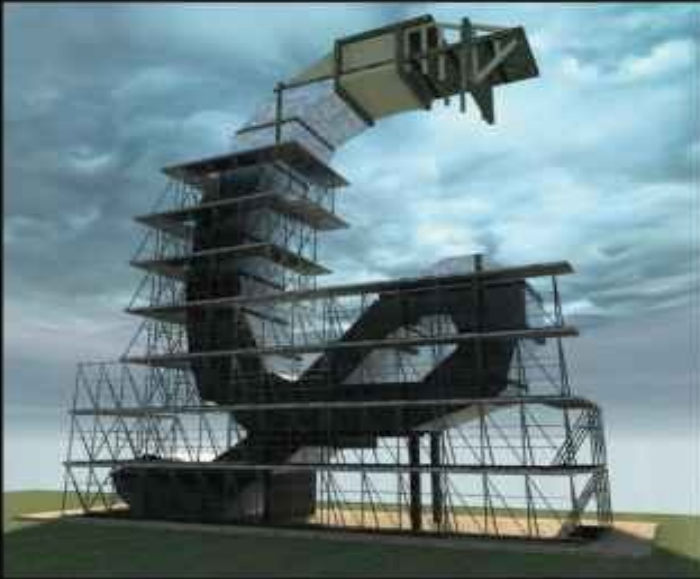


Advanced Language Construction



Mark Rosenfelder

Advanced Language Construction

by Mark Rosenfelder

≡ ⓘ YONAGU BOOKS

www.yonagu.com • Chicago • 2012

© 2012 by Mark Rosenfelder.

NW Caucasian section © 2012 by Steven Foley.

All rights reserved, including the right to reproduce this book or portions thereof in any form whatsoever, except for review purposes.

Kindle edition 2012

Contents

Introduction	7	
Writing a grammar	13	
Basic outline	13	
Plan of attack	15	
Creating paradigms	18	
Placeholders vs. filling out	19	
Wordcrafting on the go	20	
Am I done yet?	22	
Is it complicated enough?	24	
Is it simple enough?	29	
Is it weird enough?	30	
Sample texts	31	
Glosses	32	
A taste of logic	36	
Propositional logic	36	
Predicate calculus	41	
Loglangs	48	
Logographic writing	56	
Ideas	57	
Origins	60	
Pictures	61	
Determiners	64	
Stylization	65	
Cross-language adaptations	67	
On to syllabaries	69	
Dictionaries	71	
Further reading	73	
Pidgins and creoles	74	
Pidgins	74	
Creoles	77	
Origins	78	
Tok Pisin	79	
Mixed languages	81	
None of the above	82	
Further reading	83	
Life cycles	84	
Proto-World	84	
Language acquisition	86	
Bilingualism	94	
Ongoing sound change	100	

Long-term cycles	105	
Language death	106	
<i>Studies in morphosyntax</i>		109
Latin case	109	
Alignment	115	
Gender	124	
Aspect	134	
Valence	140	
Modality	146	
Mandarin syntax	156	
Polysynthesis	167	
The NW Caucasian verb		184
<i>Sign language</i>	204	
Characteristics	206	
ASL	207	
Further reading	217	
<i>Putting it all together</i>		218
Setting	218	
Overall plan	219	
Borrowings	220	
Religion	228	
Slang	231	
Nonhumans	235	
Triliteral roots	240	
Using advanced linguistics		244
In a novel	250	
<i>Tools</i>	254	
Gen	254	
SCA ²	260	
<i>Index</i>	268	

Introduction

Often when I finish an introduction to a new subject, I want to *keep going*. It's possible to read another intro, or dive into specialist subjects, but what I'd really like is Volume Two.

That's what this book is: it's a sequel to the *Language Construction Kit*, and the idea is to learn more about CONLANGING (language creation) and more about languages.

What can you expect?

- I've gone into more detail on the process of **creating a grammar**: how to organize it, what to put in it, how to write glosses, and so on.
- We'll cover **new topics** only barely touched in the first volume: logic, sign languages, pidgins and creoles, logographic systems.
- There's a chapter on the **life cycles** of languages: where they come from, how they interact, how they die.
- There's a beefy section on **morphosyntax**, which goes into much more detail on features that only got a few pages in the *LCK*: case, gender, alignment, aspect, valence, modality, polysynthesis. Also included are case studies of Latin, Mandarin, and the Northwest Caucasian languages.
- I'll show what you can do with a fully worked out **conworld**, covering topics such as borrowing, slang, and how much of your conlang you can show off in a novel.
- There's two new **tools** to explore: a vocabulary generation program and an updated Source Change Applier.

What if I didn't read Vol. 1?

You didn't read the *LCK*? Well, it's not too late! I've tried to make the book readable even if it's your first look at linguistics, but I'm not going to repeat the basics from the first volume.

You can get a lot of the basics from the online *LCK*; see the Web Resources section below.

Typographical conventions

These are the same as in the *LCK*.

I've put technical terms in SMALL CAPS. This tells you two things:

- I didn't make the term up, so you can safely use it in your grammar.
- You can get more information by Googling. (If you get a choice between (say) Linguistic List and Wikipedia, look at the former. Wikipedia isn't bad at basic linguistic terms, but it can be wrong or misleading.)

Italics are used when I'm discussing a word or phrase. If it's not in English I'll provide a gloss in 'single quotes'.

For the conventions on **glosses** see p. 32.

Referring to phonetics:

boldface refers to a letter, or informally to a sound

// slashes indicate a phonemic representation

[] brackets are used for phonetic representations

Sidebars

Sidebars give tips, warnings, or interesting facts, or provide pointers to additional reading.

Greek and Cyrillic

Aspiring linguists should already know the **Greek** alphabet, at least if they were paying attention in math class. For reference, here's the alphabet with the classical (5th century BC) phonetic values given by W. Sidney Allen in *Vox Graeca*:

Αα	Ββ	Γγ	Δδ	Εε	Ζζ	Ηη	Θθ	Ιι	Κκ	Λλ	Μμ
a	b	g	d	ε	zd	ε:	t ^h	i	k	l	m
Νν	Ξξ	Οο	Ππ	Ρρ	Σσ	Ττ	Υυ	Φφ	Χχ	Ψψ	Ωω
n	ks	o	p	r	s	t	y	p ^h	k ^h	ps	o:

σ appears as ç at the end of a word.

For the purposes of this book no harm will be done if you pronounce θ φ χ with their post-classical fricative values [θ f x]. But Πλάτων would look at you funny.

And if you know the Greek alphabet, there's little excuse for not reading **Cyrillic**. Here are the basic Russian values:

Аа	Бб	Вв	Гг	Дд	Ее	Жж	Зз	Ии	Йй	Кк
a	b	v	g	d	jε	ʒ	z	i	j	k
Лл	Мм	Нн	Оо	Пп	Рр	Сс	Тт	Уу	Фф	Хх
l	m	n	o	p	r	s	t	u	f	x
Цц	Чч	Шш	Щщ	Ъъ	Ыы	Ьь	Ээ	Юю	Яя	
ts	tʃ	ʃ	ʃ:		i		ε	ju	ja	

Russian actually has *two* series of consonants, regular ('hard') and palatalized ('soft'), and these are marked not by modifying the consonants but by changing the normal vowels а э ы о у to я е и ё ю. That is, да is /da/ and дя is /dʲa/. The second series is also used for the diphthongs j + V, as in союз 'union'.

These rules sometimes aren't enough, and that's where the unlabelled letters come in. The soft sign ь marks the previous consonant as palatalized when there's no following vowel, as in читать 'to read'. The hard sign ь similarly marks a consonant as unpalatalized; it's rarely necessary except in foreign words, such as Нью-Йорк 'New York'.

Thanks to palatalization, transliterations of Russian are either inaccurate or ugly; thus my preference for citing forms in Cyrillic.

There is of course more to both Greek and Russian pronunciation. These are transliterations, not language courses.

Web resources

There are a lot of great resources on the web. But URLs rot quickly, so instead of listing sites in the book, I'll list a single permanent URL that will be kept up-to-date with a list of links:

<http://www.zompist.com/resources/>

I mention a number of **books** in the text; these are not only great further reading, but are my major sources.

Acknowledgements

Much appreciation to Steven Foley and Drew Bennett for providing detailed information on polysynthesis; to Alex Fink and Sai for the UNLWS example; to Daniel von Brighoff for lending books and expertise, and to John Cowan for reviewing the logic chapter.

Many thanks to Daniel von Brighoff, Drew Bennett, Hannah Griffith, Anthony Duncan, Phillip Krohn, Jan Strasser, Benjamin Buckley, Matthew Pennington, Quinn Albaugh, Alex Fink, and Sai for reading the whole book, making it better, stronger, faster, but not harder. And thanks to the entire ZBB for support and useful discussions.

Also thanks to Aaron Toivo for suggestions, including a brilliant idea for the title that, perhaps foolishly, I didn't use: *The Language Construction Kaboodle*. And thanks to my wife Lida, for her patience and support, as well as checking over the Spanish bits.

Mark Rosenfelder

August 2012

Abbreviations

Here are the abbreviations I've used in glosses in this book.

>	acts upon (subject > object)	decl	declarative
*	(words) reconstructed	deduc	deductive
	(sentences) bad syntax	def	definite
?	(sentences) dubious syntax	dim	diminutive
∅	null morpheme, nothing	distrib	distributive
0	unmarked for person	dur	durative
1p	first person plural	dyn	dynamic
1pi	inclusive we (1p + listener)	emph	emphatic
1px	exclusive we	erg	ergative
1s	first person singular	exper	experiential
2p	second person plural	f	feminine
2s	second person singular	fin	finite
3p	third person plural	fut	future
3s	third person singular	gen	genitive
3sf	3s feminine	hab	habitual
3sm	3s masculine	hor	horizon of interest
3sn	3s neuter	hsy	hearsay
3x	unspecified actor	imper	imperative
4	obviative	impfv	imperfective
abl	ablative	incep	inceptive
abs	absolutive	include	inclusive
acc	accusative	indic	indicative
adv	adverb	inf	infinitive
adver	adversative	inst	instrumental
aff	affirmative	intr	intransitive
agr	agreement particle	irr	irrealis
agt	agentive	loc	locative
ant	anterior converb	lparen	left parenthesis
antip	antipassive	m	masculine
aor	aorist	masd	masdar
applic	applicative	MW	measure word
art	article	n	neuter
augm	augmentative	neg	negative
ben	benefactive	nfin	nonfinite
caus	causative	nh	nonhuman
com	comitative	nom	nominative
comp	complementizer	NP	noun phrase
compl	completive	obl	oblique
con	connective vowel	opt	optative
cond	conditional	part	participle
conj	conjunct	past	past tense
cont	continuous	perf	perfect
		pfv	perfective
pind	past indefinite	real	realis
pl	plural	refl	reflexive
plu	pluperfect	rel	relativizer

poss	possessive	retro	retrospective
pot	potential	rparen	right parenthesis
PP	prepositional phrase	s	singular
pred	predicative	sim	simultaneous coverb
prep	prepositional	stat	stative
pres	present tense	sub	subordinator
pret	preterit	subj	subjunctive
prog	progressive	topic	topic
prosp	prospective	trans	transitive
pV	preverb	voc	vocative
q	question	VP	verb phrase

Writing a grammar

Basic outline

Beginning a novel, you have to face the horror of staring at a blank page. It's easier with a conlang you can start by writing an outline! Then you can stare at a blank outline instead.

Here's the overall outline I start with:

- Introduction

- Phonology

- Morphology

- Derivational morphology

- Syntax

- Semantic fields and pragmatics

- Writing system

- Examples

- Lexicon

If you're not used to outlining, the idea is to state your topics and their order before you actually write anything. You don't write straight ahead from the first sentence of the introduction all the way to the end words in Ž. You can work on topics in any order; the outline makes sure they're in the right place and you don't forget anything.

When you think of a new topic, add it to the outline; you don't have to fill it out immediately. Topics can have subtopics, to any level you like.

For instance, you could go add subtopics to Phonology right now:

- Phonology

 - Consonants

 - Vowels

 - Stress

 - Phonotactics

Start adding text to the topics, in any order. You could start with a list of vowels (you can make a nice table later):

- Vowels

 - i e ε α o u

Any modern word processor, like Word, will have useful facilities to work with outlines. E.g. you can move entire topics around, or view just the titles of the outlines without the text.

You may find it helpful to add a symbol so you know what hasn't been filled in yet. I use STD or \$\$\$\$. Then I can jump quickly to that section by searching for this text.

Some of the sections may not make sense for a given language, or will logically appear in a different

place. E.g. if you have an alphabet, it's convenient to treat that under Phonology; while if you have an isolating language, you may have no morphology at all beyond compounding.

I am simple caveman, not know 'computer'

You can work on paper if you prefer— that's how I did Verdurian. Just expect to go through multiple drafts.

If you use a binder and loose pages, you can easily replace just a section of the grammar. Start new sections on a new page, and keep everything about a language together— avoid having your notes in five different piles or notebooks.

You can keep a dictionary in alphabetical order by using two columns. Put the words you have so far in the left column only:

bau	quiet
beda	sister
bi:	white
bori	island
buku	deer

Now, as you add new words, put them in the right column:

bau	quiet		
beda	sister	ben	under
bi:	white		
bori	island	bo	one
buku	deer		

When the page starts to get unreadable, it's time to make a new edition: copy out all the words, again only into the left column.

Index cards work too, with less rewriting but also less portability.

Plan of attack

I work on a grammar iteratively, going back and forth between sections. But my overall progress usually looks something like this:

- Put the words you have so far (perhaps from a map, perhaps from the SCA) into the **Lexicon**.
- From those words, create tables of consonants and vowels, under **Phonology**. *Always keep your phonological inventory up to date*; it guides the word creation process.
- Start the **derivational morphology** section. You'll need this even for a naming language^[1], as it's very useful to be able to form terms like "of NAME", "NAME person", "NAME speech", "NAME place". Plus it's a good habit to create derived words as you add lexical entries (*war* → *warlike*, *warrior*, *make war*).
- Create the basic **morphological paradigms** for verbs, nouns, and pronouns. At the very least you'll want the present tense forms, plurals, and a complete set of pronouns.
- Start listing your **adpositions**, or their equivalent. It's useful to be able to form expressions like "at NAME", "from NAME", "near NAME", and so on. Plus it's likely to greatly multiply your stock of verbs: e.g. *go* gives you *go to*, *go back*, *go away*, *go on behalf of*, *go with*

go near, etc.

- Start the syntax section by deciding on basic **NP order**. That is, what order do the elements go in?
 - articles*
 - numerals*
 - demonstratives*
 - adjectives*
 - quantifiers*
 - prepositional phrases*
 - relative clauses*
 - the noun itself*

Create examples and the necessary supporting words for these.

- It's useful to have tables of demonstratives, numbers, quantifiers, and prepositions, both for reference and to help you think of them as a system rather than just imitating English.
- Decide on basic **sentence order**— SOV etc. (However, that should normally *precede* NP order in the Syntax section.) Write some sample sentences, things like *The lawyer read the manuscript* and *The girl gave the book to the duke*.
- Within each section, place the basics first: the simplest forms before the insular compound forms; simple interrogatives before subordinate irrealis clauses. If you have some feature which affects a huge part of the grammar, like Old Skourene's trilateral roots (p. 235) or Elkarîl's oddball case analysis (p. 236), discuss that as early as possible.
- Now comes a long period of *filling out details*. You can take two basic approaches— or alternate between them.
 - Go through the sections of the grammar, thinking how you're going to approach each item. Work through the *LCK* and this book for ideas.
 - Work on your examples, and as you come to things you haven't said how to do, fill out the appropriate section of the grammar. E.g. to translate *The king decided to execute the man who slept with his wife*, you might have to work out the past tense, or auxiliary verbs, or relative clauses.
- At any point, you can take some time to just create words. I often use the wordlists at the end of the *LCK* (p. 260), filling out the Swadesh list or the geographical terms.
- What if you need to make **changes**? Well, it happens. Just do it: make the changes, then examine existing words and samples to get them up to date.
- At some point the outline is pretty much filled out. Are you done? The best way to finish out is to work on your **sample texts**— these days I make sure I have at least three. You'll almost always run across a few constructions you need to work out. Plus you'll have sample texts!
- Some topics should only be addressed if you need them. If you're going to write a novel in this setting, it'll be very useful to work out the calendar, common expressions, and how names and titles work. If you plan to write much text in the language, think about pragmatic particles, slang, and swearing. If this is your major language, add sections on allophony, dialects, and class variation.
- Once I have a fairly good grammar and three sample texts, the language is beginning

feel done, but I probably only have about 500 words. That's a lot for a naming language, but ~~means that translating almost any text will require word creation. So create more words, till you have a thousand or so. Work through wordlists, or just translate more texts.~~

Ultra-modingo-conlangs

What's that? You want a conlang for the ages, which will awe the conlanging boards and allow you to be the one to write Volume Three? All right, take on these tasks:

A language learning textbook and at least an entire short story (as I've done for Verdurian)

Substantial spoken recordings, with more than one speaker

Descriptions of the historical stages of the language, every 200 years or so

Meaty descriptions of the major dialects, with extended texts, and maps of isoglosses (regions where we find particular realizations of phonemes, or particular cognates)

A syntax section covering dozens of constructions, following your favorite syntactic theory

At least three places where the pragmatics differs interestingly from English

A lexicon where at least 75% of the entries are not single-word translations from English, and where every word has an etymology, with semantic changes

The lexicon gives not only meanings but pragmatic complications, register differences, and historical attestations

Creating paradigms

I work out the morphology pretty early, because without it I can't create sample sentences. You can leave gaps, but it's hard to (say) introduce a whole new dimension of verbal conjugation late in the process.

The key moment in creating a paradigm is not deciding on the affixes, but creating the structure of the table. So if you create a blank table

<i>person</i>	<i>sing.</i>	<i>pl.</i>
1		
2		
3		

you've already decided that your verbs are conjugated by person and number—and already eliminated interesting alternatives like obviative, dual, gender, and politeness forms!

Similarly you can easily create a present tense paradigm, then past and future, and not even realize that you never considered aspect, modals, or irrealis forms.

So, take a moment before filling out the table to think about whether it has the features you really want. (You can add more dimensions later; but if you do, don't forget to check your sample sentences in case they need updating.)

If you look at an actual paradigm, like the present tense of French *finir* 'finish'—

<i>person</i>	<i>sing.</i>	<i>pl.</i>
1	fin-is	fin-issons
2	fin-is	fin-issez

you may wonder where all that juicy variation comes from. How do you know how different to make the endings, or how many identical endings speakers will put up with?

- If you have a parent language, run the entire paradigm through the SCA (p. 260). Then try to simplify the output with analogy.
- For a fusional language where you don't have the parent worked out, *simulate* the above process: start with a regular, agglutinative system, then mangle it.
- Fusional paradigms are often *partially* regular. So it may be fusional *except* in two of the forms.

It may be helpful to think about where that beautiful French paradigm actually came from.

- Indo-European originally marked the three persons with final *-m*, *-s*, *-t*. 5000 years of sound change have played havoc with this, but we still see the 3rd person *-t*, as well as the *-s* in the 2s and the *-m-* (changed to *-n-*) in the 1p. (They're all silent in French, but maintained in the orthography.)
- Indo-European however didn't come up with a consistent way to mark the plural; a different method was used in each person (and to boot, in each subfamily). For another example of multiple pluralizers see Quechua, p. 171.
- The *-i* is really part of the root—it appears in every form of *finir*. The equivalent *-i* in other conjugations is less stable (e.g. the *-e-* in *parler* 'speak'), so it's convenient to treat it as part of the suffix.
- *-iss* isn't really a plural marker; it's the *-i* from the root plus the Latin inchoative *-sc-* (see p. 135).

I like to keep the Morphology section focused on the paradigms, leaving their usage to the Syntax section. That's for two reasons:

- It keeps the Morphology pages compact, making them a better reference for the paradigms.
- The usage section can then address compound tenses, auxiliaries, and other issues that don't really go under Morphology.

But you can discuss the uses of the paradigms as they come up, if you prefer. In that case, a chart of just the paradigms may be useful (i.e., a few pages containing compact morphological tables, so it's easy to look up forms).

For complicated paradigms, as for Old Skourene (p. 240), I've created Javascript conjugation utilities. If you can code, these can keep you from making mistakes in your own conlang.

Placeholders vs. filling out

If you're aiming at a grammar like mine, it's apt to be 25+ pages of dry linguistic prose. Don't be intimidated by the task of generating all that text. Start with placeholders, like this:

Questions: auxiliary verb **pol**

Assuming you've figured out how auxiliaries actually work, that's all you need to use the language. In the final Munkhâshi grammar, I expanded this as follows:

Questions use a combination of topicalization and an auxiliary; **pol** ‘do’ must be used if no other is present. The subject is fronted together with the auxiliary:

Wowal gotalh threwap tujno?

do.A.past ktuvok eat.A.past iliu

Did the ktuvok eat the iliu?

Gpuki tutujno matâ?

can.E-pl pl-iliu swim.E-pl

Can iliu swim?

The question is **answered** with appropriate forms of the auxiliary: **Wothôl** ‘Yes, B is going’; **Potôrul** ‘No, D isn’t going.’

It’s not just a matter of writing full sentences. Trying to explain the procedure, you’ll find you have work out minor details. In this case: what if there’s another auxiliary; how is the question answered; what about negative questions (not shown).

It’s work to create sample sentences and glosses, but every sentence you write is another chance to develop the vocabulary and add new points to the language.

Wordcrafting on the go

As you work on the grammar you’ll be inventing words; never create one without adding it to the lexicon, in alphabetical order. Not only does this ensure they don’t get lost, but it keeps you from accidentally creating homophones. (A few homophones are fine, especially if they’re not likely to occur often. But it’s easy to create too many, especially if you use a vocabulary generator.) Plus, it’s a lot of work to generate a lexicon, so every bit you do gets you closer to the finish!

E.g. the Dhekhname word for *swim* was entered into the lexicon like this:

math	v	swim, float [<i>mat</i>]
-------------	---	----------------------------

I always use a table format, which looks neater than straight text. If there are morphological peculiarities (such as the out-of-control plurals in Xurnese), I indicate these in a column just after the word itself:

púsaup	pusú	<i>n</i>	poor bastard [<i>poukuvi</i> ‘fallen’]
puš	pauč	<i>n</i>	stomach, abdomen [<i>puč</i>]
puxamu	puxamú	<i>n</i>	return [<i>poudixamou</i>]
pwes	pwesi	<i>n</i>	pebble, stone [<i>puvik</i> dim. of ‘stone’]

(Some languages have a morphology that just spits on alphabetical order— e.g. Old Skourene *agaš* ‘beloved’, *egušet* ‘romance’, *gšit* ‘affair’, and *iggšet* ‘loving’ are all formed from one root. So the lexicon is sorted by roots, and all these words are entered under *gašt*- ‘love’.)

It’s a good habit to provide a **part of speech** column. This provides a place for morphological data (e.g. gender of nouns, conjugation class for verbs), it disambiguates glosses (e.g. ‘a bear’ vs. ‘bear’), and it allows searches— e.g. you can look for all your prepositions.

Another good habit is to provide **multiple glosses**. Fight the tendency to make every word a one-for-one equivalent of one English word. This makes your language more naturalistic, and can save time later when you find you need the other word.

Extra credit if you take the time to work out some quick **derivations**. E.g. *swim* could generate words for *swim* (*n*), *swimmer*, *swimming hole*. Extra extra credit if some of the derivations aren't already in English. E.g. *swim-thing* might be the word for *fish*; *swim* + *diminutive* might be *bathe*.

I hate to create a word without an **etymology**. Dhekhname is created mostly from Munkhâshi using the SCA, so to invent *math* I actually created *mat*, added it to the Munkhâshi lexicon, and ran it through the SCA. Often I'll borrow the word instead (p. 220), or derive it as a compound.

Words usually don't retain a single meaning for millennia on end— you should often take the opportunity to modify the meaning of an inherited or borrowed word (p. 226).

How do you **look up a word** when you need it? Well, you're doing this on the computer, right? Use the search function. If it's a common word, you can save time by placing the cursor at the beginning of the lexicon, or just keep your lexicon in a separate file.

An alternative is to include a separate English-to-Conlang lexicon. That's not a bad thing to have, but it's a huge hassle to maintain, and it makes it all too easy to create ciphers of English— e.g. you create a word for *can* and later when you want to translate *ability* you create a different word just because *ability* doesn't yet have an entry. So it's best to create such a lexicon when the language is pretty much done.

Am I done yet?

You read the *LCK*, so you know you should have a *Syntax* section, and it contains the single word "VSO". What else goes there?

Here's a checklist, not at all exhaustive, of things that you should consider putting in the grammar somewhere.

- The basics: sentence and NP order; questions; negatives; relative clauses.
- Can people violate the basic order— for topicalization, for emphasis, for passives, or just as an afterthought?
- How do verb + verb combinations work? This includes auxiliaries (*You may go*) as well as ordinary verbs (*I advise you to go*).
- Where do time and place clauses live? You'll probably have single words (*never*), NPs (*last week*), and clauses (*when Oblivion freezes over*).
- Existentials (*There's a Chinese place near here*) are often a special construction.
- How do you handle sentential arguments? These can be subjects (*[That people still read Nietzsche] offends me*) or objects (*Holmes deduced [that the criminal wore a tartan]*).
- Make sure your relativization scheme clearly handles the four basic combinations of transitivity:

sentence	subclause	example
subject	subject	<i>The man [who caught the fish] is here</i>
subject	object	<i>The fish [the man caught] is tasty</i>
object	subject	<i>I hate the man [who caught the fish]</i>
object	object	<i>I ate the fish [the man caught]</i>

Pay attention to cases (which case is *who* in?) and to word order (the subclause omits arguments, so it may complicate your basic sentence order).

- How do you translate *if* statements? Are deductions (*If that's Camelot, we're almost home*) handled differently from counterfactuals (*If grandma had wheels, she'd be a wagon*)? Conditionals are a playground for seeing how your conlang handles not-quite-real events.
- How do you handle causatives? (*I made her go, I advised her to carry a gun*). These tend to play havoc with case systems as the caused person is the object of the main clause and the subject of the subclause.
- Pronouns may be an exception to word order rules, as in French *Je le lui ai donné* (*I gave it to him*) vs. *J'ai donné le livre à mon ami* (*I gave the book to my friend*).
- You worked out the numbers from 1 to 10, great! How do you form larger numbers? ordinals, fractions? How about basic mathematics?
- Are there restrictions on relativization? Think about questions like these:
Where is the hat [I believe [the cat wears ___]]?
When was the day [the cat bought the hat on ___]?
Who did [John said ["I'll kill ___!"]]?
I remember the summer [we visited Rio last ___].
The man [my sister dated the brother of ___] was a crook.
- How do you form comparatives? (See the next section.)
- How do NP + NP combinations work? You'll need these for titles (*King Alric*), geographic names (*Lake Van*), and brands (*Yonagu Books*), and there are alternatives to English's concatenation method.
- Are there ways to indicate that a referent, or a relative clause, refers to something known to exist? Compare *I met a queen* vs. *I met the queen*. In English *I'm looking for a man with one arm* is ambiguous as to whether I have a specific man in mind, but other languages differ.
- You can nominalize a verb (*know* → *knowledge*); how do you nominalize a VP? Note the combination of cases and prepositions in *John's knowledge of Linux*.

Is it complicated enough?

You may be trying for a simplified language—or you're just in a hurry to get done. But a hallmark of natural languages is their almost fractal complexity. There's always another exception or complication, and linguists can write entire dissertations on a single word.

Complexities may occur to you if you just think hard about a feature. Say you're thinking about comparatives: you work out how to say *bigger than a mammoth*. Revolve the concept of comparison around in your head—does your method work on these cases?

- superlatives (*biggest of all*); note that speakers may turn absolutes into intensives (*fortissimo* 'strongest' often means just 'very strong')
- degrees of comparison (e.g. mathematics uses \gg for *is very much greater than*)
- equalities (*as big as a mammoth*; note the difference from the morphological comparative)

bigger)

- negatives (*no bigger than a fly*)
- examples with and without a comparison class (*a better mousetrap; a mousetrap better than Roger's*— hey, the word order changed!)
- comparisons of adverbs (*more slowly*) or verbs (*he cried more than he laughed*)

You can't always think of such complexities just staring at the computer. Alternatives include looking at other people's grammars, and waiting till interesting cases come up in sample texts.

Sometimes an idea that didn't make it into the morphology may pop up elsewhere. E.g. French doesn't have evidentials, but it can use the conditional as one: *il aurait allé* can be used for 'he supposed went'. English doesn't have a topic particle, but clefting is a substitute: *what I'm looking for is a cheap bicycle*.

Another source of complication is to think about variations of dialect or register. Come up with three ways to solve the problem and assign one to the yokels from Nowheresville and another to colloquial speech. If you've derived your language from a parent, the newer language may have innovated a new method but kept the parent's method in formal written language.

Ten quirky constructions

Languages are full of minor constructions with their own odd syntax; here's a sampling. You don't have to address these in particular; the point is that once you start looking you'll find more and more (The asterisk indicates sentences that aren't acceptable.[\[3\]](#))

Let alone

I wouldn't live in Vyat, let alone Verduria.

She won't pet the dragons, let alone clean up their dung.

This may seem straightforward, but what type of constituent is the 'let alone' phrase? And where does it come from? We have *let (NP) alone* in other contexts, but can't move the NP:

Let the boy alone!

**Let alone the boy!*

What, dative + VP

What, me worry?

What, him get elected?

It looks like something got left out, but what? And if it's a deletion, why is it allowed only after *What*?

**How, me worry?*

Damn them!

Fuck you.

Damn those robots.

Yes, you can do syntax on profanity. The oddity here is that the object isn't reflexive, as in *Fuck yourself!* Maybe it's not an imperative but a wish, perhaps an abbreviation for *If only someone would fuck you*. But then why can't we make a similar abbreviation for *If only someone would kill you*?

Do so

Itep cheated on the test, and Deru did so too.

Do so is interesting because it's a verbal ANAPHOR— just as a pronoun stands for an NP, *do so* stands for a VP.

In Verdurian I created a demonstrative verbal anaphor: *fassec* means *do that*:

Vulre dy žusru soa card'ä er faššao.

want-past-3s that sell-past-1s the.f.acc sword.f.acc and do.that.past-1s

He wanted me to sell the sword and I did it.

Quechua has the verbal anaphor *nay* which stands for a verb you can't think of the moment— *whatchamacallit* for nouns:

Wallpata narankichu?

chicken-acc whatsit-past-2s-q

Did you do that thing to the chicken?

Hard to like

My grandmother is hard to like.

It's hard to like my grandmother.

These have been derived transformationally from

[(For someone) to like my grandmother] is hard

though note that not all adjectives work:

[(For someone) to like my grandmother] is outrageous

→ **My grandmother is outrageous to like.*

The surface form is identical to *The elf is eager to please*, but the semantics differ: my grandmother is the underlying object; the elf is the underlying subject. Noam Chomsky used such sentences to show that the surface structure of sentences isn't enough to determine their meanings.

While we're at it, this construction is an example of an ADJECTIVAL PHRASE, something you might not expect to see if you thought of adjectives as nothing but modifiers. Compare also *afraid to go*, *capable of flight*, *important to know*, *lucky to be alive*, *glad you came*, *new to the city*.

As well as

John put books as well as records in the closet.

Chris played Deus Ex on the PC as well as on Xbox.

We've had pizza yesterday as well as today.

As well as (and similar expressions: *in addition to*, *rather than*, *instead of*) look like conjunctions. But, curiously, they don't play well with VPs or entire sentences:

?*It rained as well as snowed.*

**John looks like Justin Bieber as well as owns a bank.*

**We had pizza as well as Julie did a dance.*

Each other

The angels sang to each other.

sample content of Advanced Language Construction

- [download online Ben Sidran: A Life in the Music](#)
- [More Die Of Heartbreak pdf, azw \(kindle\)](#)
- [click The One I Left Behind book](#)
- [download Warrior's Curse \(The Warrior Chronicles, Book 5\) pdf, azw \(kindle\)](#)
- [read Think Big: Unleashing Your Potential for Excellence pdf, azw \(kindle\)](#)

- <http://nexson.arzamaszev.com/library/How-to-Rescue-a-Dead-Princess.pdf>
- <http://transtrade.cz/?ebooks/Jurassic-Park.pdf>
- <http://pittiger.com/lib/The-One-I-Left-Behind.pdf>
- <http://test.markblaustein.com/library/One-Constant-Star--Star-Trek--The-Lost-Era--Book-11-.pdf>
- <http://damianfoster.com/books/The-Survivor--Bill-Clinton-in-the-White-House.pdf>