

Making Dictionaries that Support Bible Translation Projects

A new reference manual for Nigerian Bible Translation Projects (2012)
David Rowbory — david_rowbory@wycliffe.org

An initial survey:

- Does your language have any dictionaries?
- What benefits would a dictionary bring to your language development?
- Do you own any dictionaries?
- What do you use dictionaries for?
- What dictionaries do you look at? Why?

What is a dictionary? Who should make one? Why?

Especially since the invention of printing, various different kinds of reference book have been developed, to allow people to share their knowledge.

- A **dictionary** puts in writing what a community knows about their language.
- An **encyclopædia** puts in writing what a community knows about the world.

Everyone carries their own personal dictionary in their head. People who know several languages have several different dictionaries in their heads perhaps with links between them. Dictionaries help both readers and writers to use their language to communicate powerfully.

Outsiders are interested in dictionaries of other languages for various reasons, but this course is for pioneer mother-tongue linguists. We aim to help native speakers build practical dictionaries that help others read and write their language better. Many people who compile dictionaries are involved in Bible translation and need to study the meaning of words to find the best way of translating the original texts. A diligent translator who works on a dictionary should discover that the skills and experience of compiling a good dictionary benefits his work as a translator. The rich store of information in a dictionary will help other translators, advisors and consultants draw on the knowledge of a whole community to produce a Bible that will build up that whole community.

A good dictionary is not just a large book that gives a community a sense of prestige or props open a door, but it is a living, growing thing. A dictionary is the fruit of a community's communication and the seed for ongoing language development.

Different kinds of language reference book

Apart from **dictionaries** and **encyclopædias**, you might also hear about glossaries, word lists, lexicons, thesauri (or thesauruses), grammars, primers and a few others.

- A **glossary** briefly explains technical or unfamiliar words using the same language. For example Bibles or textbooks use glossaries to explain **key terms**. (This list is a non-alphabetic glossary.) Most large glossaries organise words alphabetically (A-Z) to make words easier to find.
- A **word-list** or **lexicon** gives a very short 'gloss' to explain the meaning of a word in another language. For example a Greek-English lexicon gives quick English explanations of Greek words.
- A **thesaurus** or classified dictionary arranges words not alphabetically, but according to meaning, so that similar or related words are listed together.

- A **grammar** either describes how a language organises words into phrases and sentences, or teaches an outsider how the language works.
- A **primer** helps people learn to read and write a language they already speak.

More about thesauri: A thesaurus is a prompt to help writers find the best words to use to convey the meaning they intend. The word came into English from Latin from Greek. (It originally meant a storehouse or treasure.) Roget published a 'thesaurus' of English Words and Phrases in 1852.

Why are you making a dictionary?

Who is your audience? Do you have several different audiences in mind? Why would they use the dictionary? Consider your answers to these questions then think about how each of your audiences might access it. Would it be best to produce several different dictionaries to serve each group better?

Various purposes that people sometimes give:

- **Prestige:** A large published dictionary will give our language prestige, showing that it is a real language like Hausa, Yoruba, Igbo and English which have their dictionaries.
- **Preservation:** Younger people and those in large towns are forgetting some of our words. The dictionary will remind people of the resources we have and encourage them to use our language rather than Hausa, English etc. Outsiders may fear that a language is dying and want to record some of the words so as to preserve a historical record.
- **Research:** Outsiders want to record the words of the language so they can compare with neighbouring languages and do other research on languages in Africa.
- **Checking Spelling:** A dictionary tells us the right way to spell a word which is especially important for new writers.
- **Faster Translation:** A dictionary will enable people to translate the Bible easily because then anyone can look up an English word and then find the right word in the language. (Note: this is not a good idea.)
- Perhaps you will hear more.

Some of the above reasons are good in parts, but all can be misleading. There are other, better ways a dictionary can benefit a community and writers. Real prestige for a language community is not obtained by producing a book which is never opened, but by using and developing their language in spoken and written form, for all important areas of life. A dictionary may help preserve a language, but if the language dies or changes then a historical record will not help anyone communicate. Research done by outsiders may or may not benefit the community because the objectives are very different. Correct spelling can be reinforced by a dictionary, but to look up a word you need to know its approximate spelling already. Finally translation is not a matter of replacing English words with those of another language. However, advisors and translators may avoid wasting time if they have access to a good dictionary.

Remember, every different audience has different needs in a dictionary. Undoubtedly the general community is the first audience that you will think of, but the translation team can also use the dictionary and benefit from the process.

Why and how might a translator use a dictionary?

Almost no Bible translation projects think about making a dictionary until after they finish their work. Even fewer translators realise that they can use a dictionary to make a better Bible translation.

It has been common for people to assume that a mother-tongue translator doesn't need a dictionary; he stores all the knowledge of his language in his head. This is in contrast to a foreign translator who may often need to refer to a (bilingual) dictionary to find the right words to use, because they have not grown up speaking the language. However, I believe this assumption at least needs to be questioned. A mother-tongue translator will not rely on a dictionary in the same way that foreign translators might, but depending on the situation there may be some good reasons to make use of a dictionary. Since the audience of the dictionary is very important for the writing of definitions, it is important to consider how mother-tongue translators (of the Bible and other materials) may refer to the dictionary.

Here are three examples that come to my mind. (There may be more and some of these may not always be relevant.)

Dialect differences: It is sometimes hard to remember the dialect variation of particular words – whether certain dialects have different pronunciations, or different meanings or senses of words. A dictionary is a good place both to record the differences when discovered and then to refer later so that all the dialect variation can be taken into account. That way it is possible that the translation team can learn from comments made by reviewers from particular dialects, rather than forgetting them.

Wider vocabulary options: Especially when writers are beginning to learn to write in their language, or for writers who are not naturally the best communicators, it can be helpful to have lists of related/similar words (a thesaurus) and a dictionary. Good definitions and example sentences let a writer check that any word is being used correctly, so that readers will understand correctly.

A reminder for those influenced by other languages: Many Bible translators and pioneer writers were first literate in other languages. Some have studied and lived for some time out of their language area. When translating anything, new translators may be so controlled by the words and the writing style of the original text that they just can't think of the appropriate word in their language, or they use words in an unnatural, stilted way. In such situations, a dictionary can be a helpful reminder of the real meaning of words in the language. Sometimes one translator may not immediately think of the best words to use, but use plain or general words. This results in a boring translation, but a large store of words in a language (a classified dictionary or thesaurus) can remind the writer of words he might have forgotten.

To help consultants help you more: Both new and experienced translators work with consultants to help make sure their translations are accurate, clear, natural and acceptable to the audience. Most consultants don't speak the language much, but with a good dictionary, they can help translators much more. A consultant needs not just glosses and but clear definitions in English (or Hausa)

Making a dictionary develops skills that Bible translators need

A dictionary can not only help translators and new writers in a language, but the people who work hard on the dictionary will gain skills that help them become better Bible translators. The process of making a good dictionary helps the **dictionary authors** in at least three ways:

- They learn how to describe the differences between similar/related words;
- They learn new words and idioms that they might not know, but which will help them use the language more powerfully;
- They recognise borrowed words

Bible translators need to think carefully about the meaning of words as used in the Bible, and then use the best ways of expressing that meaning precisely and powerfully in the language. A diligent Bible translator uses Bible dictionaries and other aids to understand the Bible well. They also need to learn to think carefully about the best words to choose from their own language. Sometimes translations can be misleading, unclear or boring because the translator has never thought carefully about his own language. There may be many different ways of saying similar things but

unless you have really worked hard with the words of your language, you cannot be sure you are using it well.

Building a dictionary should involve a large number of people from the **community**. This itself can bring benefits. It should raise awareness in the community of the language development and Bible translation programme. It shows that this is an activity which is to benefit the whole community but which requires a large number of people for it to be effective. It is therefore good to share the community’s treasure of words with the whole community as cheaply as possible, which is why electronic publication is helpful.

Some example audience needs – complete the missing boxes yourself:

Audience	Needs
Experienced writers of the language who are not literate in any other language	<ul style="list-style-type: none"> • Detailed definitions in the language • Links to related words • Large number of words
Experienced writers of the language who also read English	<ul style="list-style-type: none"> • Detailed definitions either in the language or in English • An index from English → the language
Adult speakers of the language who are literate in English but not in the language	<ul style="list-style-type: none"> • English glosses and definitions. • English index • Pictures
Adult speakers of the language who are literate only in Hausa	<ul style="list-style-type: none"> • • •
Adult speakers of the language who cannot yet read or write well in any language	<ul style="list-style-type: none"> • • •
Older children who learn English at school, speak Hausa around school and speak English in their village	<ul style="list-style-type: none"> • • •
Young children in the village who speak the language	<ul style="list-style-type: none"> • • •
A Nigerian pastor or businessman from elsewhere in northern Nigeria who lives in the language area	<ul style="list-style-type: none"> • • •

Audience	Needs
A translation advisor or consultant from Nigeria or abroad who is helping you check and improve your Bible translation	<ul style="list-style-type: none"> • • •
A linguistic advisor or consultant who is helping you develop your language	<ul style="list-style-type: none"> • • •

One good example audience statement for a dictionary for language XXX

This dictionary is firstly for XXX children learning to read and write in their own language before they learn to read another language. Secondly this should support adult speakers of XXX become confident in using their language in written form. Other dictionaries may be better suited for outsiders to learn the XXX language.

(The dictionary will probably have definitions in the XXX language and numerous pictures and example sentences to show correct writing.)

Another good example audience statement

This dictionary is firstly designed for adult speakers of XXX language who are already literate in either English or Hausa and want to learn to write their own language well. It may also help non-speakers of XXX, such as translation consultants or workers from other parts of Nigeria, understand XXX words better.

(This dictionary will have meanings given in English and Hausa, a pronunciation guide, possibly an English → XXX index and other material useful for outsiders.)

A bad example audience statement for a dictionary for language XXX

This dictionary is aimed at the widest audience: to serve everyone in every way. We have therefore included everything that anyone might want to know about the XXX language.

(This dictionary aims to do everything and so will probably do it very poorly and unevenly.)

What information should you record for a word?

The audiences that you have in mind determine what information you choose to include in a dictionary and how you complete that information. However, you can use a single WeSay dictionary to create several different dictionaries for different audiences. This means that it is best to complete as much information as possible about each entry in the dictionary. The simplest descriptions are **word lists**:

Mada	Hausa	English
ngwon	hannu	hand
...		

Word lists match one language word with the closest ‘gloss’ (or equivalent). They can be very useful for seeing historical connections between languages and for understanding the sounds used by a language. However, word lists are not satisfactory for any detailed discussion of meaning. The **part of speech** (noun/ verb/ other...) is very important and so is a **full description** of the meaning of each word, divided into different senses or uses. Here are some different ways that words can be described.

Minimum information	Other information
<ul style="list-style-type: none"> • Word • Meaning / Definition(s) • Gloss(es) • Part of speech / kind of word 	<ul style="list-style-type: none"> • Variants • Semantic Domain • Sound recording • Picture • Plurals and other forms of the word • Related words

Information for describing words: an explanation...	
Word	<p>For most words this is the simplest form – such as a singular noun.</p> <ul style="list-style-type: none"> • You must mark the tone on the word, probably by using accents such as <i>é à...</i> because two words with the same consonants and vowels but different tone are really two separate words. • You must think hard about the most helpful way to represent verbs.
Variants	<p>If different dialects have different pronunciations, include the dialect name and the way they write or say this word.</p> <ul style="list-style-type: none"> • [Riba] <i>ryá</i>
Meaning / Definition	<p>The meaning is a full explanation of each sense.</p> <ul style="list-style-type: none"> • Make sure you separate the different senses or meanings. • For a monolingual dictionary, you write the definition in the same language as the word you are explaining • For a bilingual dictionary you should write a gloss, then semicolon, then a further explanation of the exact meaning of the word, even if the gloss makes it seem obvious to you.
Gloss	<p>A gloss is a word or short phrase in another language that is close in meaning to the language word.</p> <ul style="list-style-type: none"> • If you can think of several different glosses and you are not sure which is best, separate them with semicolons ;
Part of Speech	<p>The part of speech is the type, kind or category of word and is nothing to do with its meaning; it is the way the word fits into the structure of a sentence.</p> <ul style="list-style-type: none"> • Common types are nouns, pronouns, verbs, conjunctions, ideophones. • Other kinds of words might work together with nouns or verbs, but you must investigate what categories your language has.
Plurals and other forms	<p>Each different part of speech needs special information:</p> <ul style="list-style-type: none"> • Nouns always need a plural form. • Verbs need other forms – any forms which are not completely predictable. This might be a ‘plural’ form of the verb, or possibly different tenses or aspects if the main verb word changes.
Semantic Domain	<p>The semantic domain is a way to link similar or related words together by their meaning. A word might belong in several different semantic domains, especially if it has different senses.</p> <ul style="list-style-type: none"> • The most common way to record Semantic Domains is to use the reference numbers in the DDP (Dictionary Development Process). • If you gather words in WeSay using the Semantic Domains method, the reference numbers are recorded by WeSay.
Sound recording	<p>Record the sound of the word so that a non-native speaker can recognise the word, or to help new writers/readers connect the written and spoken forms.</p> <ul style="list-style-type: none"> • This will also help you get advice in case you have a spelling mistake or are unsure how to record the tone. • You can also record example sentences.

Information for describing words: an explanation...	
Picture	Some words are most clearly explained with a picture. Some audiences (such as children or pre-literate people) may find dictionaries with many pictures are easier and more interesting to use. <ul style="list-style-type: none"> If you have the Art of Reading in the C:\ drive, WeSay gives you a list of pictures to choose from based on the gloss you have given.
Related Words or Cross Reference	It may be useful to make a list of related words – whether similar words, opposites or other relationships.

Technical note for WeSay users: If you are using WeSay, you might need an advisor to create new ‘custom fields’ for the **variants** and **other forms** (such as plurals, past tenses etc) and choose to display hidden fields such as **Cross Reference**, **Gloss** etc. If you are trying to add information in the **Dictionary Browse & Edit** view, you might need to click **Show Uncommon Fields**. For sound recordings, you may need an advisor to include the **voice** Input System for the **Word** field.

Editing: Advice for making good, helpful dictionaries

A dictionary-maker often has difficult decisions to make: what words or phrases to include, and how to describe them. If you have a clear focus on the purpose of your dictionary and if you know your audience, then you should think about how someone will use the dictionary. Make sure you write the dictionary so that it is as easy as possible to find helpful information. Keep using the dictionary yourself and test whether other people are able to use it well. If the dictionary users get confused, lost, frustrated or don’t learn anything from the dictionary then it can be improved. Be selective in the information you include, but don’t assume that everyone knows what you know. Don’t repeat yourself. Use pictures or example sentences only if they actually help describe a word. Think about whether information you gather is best included in the main part of the dictionary or whether it will be more helpful to keep it in a separate place: such as a body chart, or various other lists of names, places, cultural information, the calendar etc.

Choosing citation forms / headwords

One of the hardest decisions a dictionary-maker has to make is what form of a word should be chosen to represent all the forms of that word. This is particularly difficult for words that take a number of different prefixes. This will probably affect many Nigerian languages that have a noun class prefix, but it might also affect verbs, if the subject changes the prefix on the verb. Suffixes don’t matter so much because dictionaries are normally arranged in alphabetical order from the start of each word. You might need to test different possibilities to make sure that people can find words easily.

In the C’Lela dictionary (2003), we ignored the noun class prefix for organising the words, but always displayed the most common prefix. So, for example, **C’Lela** (the language) is listed in the **L** section of the dictionary, but is displayed as **C’Lela**, just after the word **A’Lela** (the area where people speak **C’Lela**) and before the other related words **D’Lela**, **K’Lela** and **Lelna**. Similarly **d’taare** (stone) is in the **T** section of the dictionary. It would be good to test with different audiences whether they find this helpful or confusing. You may need to change options in WeSay and possibly move to using FieldWorks to do some more complicated things like this.

One advantage of electronic dictionaries is that someone can search for a word and they might find it even if they don’t know what to do with the prefix. So for electronic dictionary publication the exact choice of citation (or reference) form is not so important.

Glosses & Definitions

A **gloss** is a word in English or Hausa that is close in meaning to the word in the language. This is what most people think of immediately.

A **definition** is a longer explanation either in English, Hausa or even the language itself, to explain the meaning of the word clearly and precisely. A definition is much harder to write but much more useful than a gloss. It is even more important to remember your audience when you write a definition, than a gloss. Two different words might have the same gloss but shouldn't have the same definition in a dictionary, because no two words ever mean **exactly** the same thing.

Examples of how the same gloss might actually be used to mean different things:

- Mada word: **ngwon** English gloss: 'hand'
 - definition: 'arm; part of the body from the shoulder to the fingertips'
- Greek word: **cheir** English gloss: 'hand'
 - definition: 'hand; part of the body including everything from wrist to fingertips'
- English word: **hand**
 - definition: 'part of the body from below the wrist sometimes including the fingers'
- Different languages might have words glossed in English as 'tree' but actually mean different things, perhaps:
 - definition A: 'a plant that grows high with a strong trunk, taller than a shrub'
 - definition B: 'a plant that bears fruit'
 - definition C: 'any plant that has a brown trunk or stem'

Why detailed definitions matter

It is very important for Bible translators to learn to write good, precise, detailed definitions of words in their language because a translator will need to think hard about many Bible terms (or **Key Biblical Terms**) and choose the best terms amongst similar possible terms. People often **wrongly** assume that you can match one word from the language exactly onto English or Hausa or another language. There are never any two words which are exactly the same, whether in the same language or between different languages. If you think two words are exactly the same in every way, then you may have forgotten some details or the full range of the ways that the words can be used.

Senses

Some words have several different **senses** where words are used in different ways, with different kinds of meaning intended. Usually there is a common meaning, then figurative or extended meanings. For example the basic sense of **brother** in English is '**male child of the same father and mother**', but there are several extended senses:

- **[Nigerian English] male relative (general)**
- **[Christian] fellow Christian believer**

In most varieties of Hausa the word **itace** can be used in two different senses:

- Sense 1: gloss: **tree**
 - definition: **tree; a tall plant, larger than a shrub, whether still growing or not**
 - *(This is the object as it is growing.)*
- Sense 2: gloss: **firewood**
 - definition: **wood from a tree that has been cut down**
 - *(This is an extended sense focussing on the purpose or utility of the wood of the tree. We know that you don't cook with a tree, but you cook with firewood.)*

There may also be quite a few other extended senses of the word. Sometimes words have a technical sense. (This is true even of the English word ‘sense’ which has a special linguistic meaning, but which otherwise could mean ‘intelligence’ or ‘hearing’ or ‘sight’ etc.) The context may make it clear what meaning is intended, but a dictionary should describe each sense beginning with the most basic sense.

Warnings: common misunderstandings

- Senses are not the same as different English glosses. So Hausa ‘**gida**’ could be represented as ‘house’ and ‘home’ but these are really the same meaning.
- Multiple senses must share some common idea, as the different senses of Hausa ‘**gida**’ (house, household, descendents) or ‘**ruwa**’ (water, rain). If not, then they are separate words (same-sound words) and need separate entries in the dictionary. An example in English would be the words ‘right’, ‘rite’ and ‘write’ which sound exactly the same but which have no relationship in meaning at all.

How distinguishing senses benefits a Bible translation:

Bible translators need to be able to think about the different senses and uses of words, in particular when it comes to **Key Biblical Terms**. Some words from the Bible have different senses and it is important to know how the different senses of words in your language can best represent each sense of the Bible word.

Show sentences / Demonstration sentences / Example sentences

Some words are hard to describe and define. A ‘show’ sentence can make things much clearer. For example:

- English ‘and’: I went to the market **and** bought rice, beans **and** flour. (≠ Hausa ‘da’)
- **Because** of the heavy rain, the road became very muddy.
- The car **crashed** into a tree.
- The tree fell down with a loud **crash!**

When should you use example sentences?

- They are especially important for: **connector words**, **ideophones** and **helper words**.
- They can show how a verb fits into a sentence.
- They are **not usually necessary** for nouns, unless a definition is difficult.
- They are a good way to demonstrate different senses. (Each sense has a different collection of example sentences.)

Good show sentences...

- ...are short, but not too short: long enough to give suitable context.
 - eg connector words probably need 2 clauses at least.
- ...help you understand the meaning or purpose of a word from the context, even if you have never seen that word before. (Try replacing the word with a space.)
 - eg: I shouted your name [-----] you didn’t hear me. [but,however]
 - eg: The girl started crying [———] she had fallen over. [because,since,for]
- ...show several different ways of using the word
- ...must be checked carefully to make sure spelling and all other details is correct
- ...are helpful for new writers and for consultants, to help Bible translations use these words naturally and powerfully in translation.

Bad show sentences...

- ...add nothing to the description.
- ...presume you already know what the word means.
- ...have spelling or grammar mistakes.

How can you get good show sentences?

- Look in your texts. Perhaps simplify some sentences and use them as examples.

In WeSay you type an Example sentence in the box marked **Example**. Press the **tab** key and you should be able to add a translation underneath.

Continued later...

Return to www.ngbible.com/tools to find more information through 2012 and 2013 to help you discover the richness of your language and build better dictionaries.

Other sections to be given in written form:

- Companions to the lectures
 - Different ways of gathering more words
 - Publishing options
 - More details on recording dialect differences
 - Idioms and phrases in the dictionary
- How to use WeSay
 - including how to Send & Receive your dictionary over the internet
 - How to change the way that WeSay works
- How to produce electronic dictionaries
- How to produce printed dictionaries:
 - Small topical dictionaries
 - Full printed dictionaries
- Historical notes on preparing dictionaries using index cards and other paper methods.
- Alphabetisation
- Producing better descriptions:
 - Describing senses using concordances and collocations
 - Connotations, emotive and associative meanings
- Subentries and cross-references
- Body charts, other annotated pictures and material for appendices