

One dictionary, one language, one team, but different locations?

Version control and file management turn chaos into quality*

Charles E. Grimes

Linguistics, RSPAS, ANU, and SIL International

chuckgrimes@bigpond.com

Abstract

The increased availability of laptop computers, along with readily available dictionary software like SHOEBOX and its direct descendant TOOLBOX, have made it feasible to work on dictionaries of minority languages as TEAMS (rather than as a scholar with one or more assistants), to work on multiple computers, even to work in multiple locations. This is on the increase as foreign scholars become more aware of the advantages and very real possibilities of training mother-tongue speakers of languages to take more active roles in compiling dictionaries. However, as there are more attempts to work successfully as a team in this way, I am getting fairly consistent reports from around the world of frustration, mangled data, and chaos.

This paper identifies some of the common problem areas, and discusses some solutions that have been tested and shown to work in typical situations. It discusses how to manage information and files when there is a team working in different locations, or on different computers. It describes some fairly simple procedures and safeguards for how to keep things from getting out of control.

1. Background experience

In the years leading up to and immediately following publication of the Multi Dictionary Formatter manual (Coward & Grimes, 1995),¹ most of the queries and requests for consultation on dictionaries that I received from around SE Asia, the Pacific, Africa, and Latin America were from individual scholars compiling dictionaries on their own, either for their own academic purposes, or with the intention of the dictionary also being usable by the local community. For most of these dictionaries, the native speakers and the local communities played an adjunct role of 'language resource', supplying data on request, rather than collaborator or co-compiler with ownership and decision-making input.

Occasionally during those earlier years, but much more frequently in recent years I have been getting requests for consultation on dictionaries being compiled by teams—for national languages, regional languages, and local vernaculars. I have been part of a team that published a dictionary of a language of wider communication (Jacob & Grimes, 2003), even though key members of the team were living in different countries. Several similar dictionaries are in process. (See references.)

In June 2005 I led a lexicography workshop in Malay for 62 participants speaking 16 languages spread all over Malaysia (15 Austronesian and one Mon-Khmer). Almost all of these languages had two or more participants, and two or more computers for their team. One language had eight native speakers making entries on four different computers at the same time.

* Paper presented at the Panel on Dictionaries, Dictionary-Making and Lexical Semantics at the 10th International Conference on Austronesian Linguistics, Puerto Princessa, Philippines, 17-20 January 2006. I am grateful for comments and suggestions on an earlier draft of this paper by Joseph Grimes.

¹ A PDF version of the MDF manual is available on CD with the purchase of the SHOEBOX computer program. It is also available for download at <http://www.sil.org/computing/shoebox/MDF.html>.

2. Problem areas

Compiling dictionaries as a TEAM highlights some problems and challenges that were not so obvious when one person was making all the decisions and doing all the work on a single computer. Problems arise when there are:

1. More than one DECISION-MAKER on the team
2. More than one LOCATION at which the team is active
3. More than one COMPUTER for compiling the results
4. More than one SYSTEM FOR NAMING COMPUTER FILES
5. More than one system for BACKING-UP FILES, or no back-up system at all

Unless the PROCESS for compiling a dictionary as a team is carefully defined and tightly followed, experience shows that chaos results. Among the teams that have approached me for help, it is fair to say that most of them did not define a process or think through some basic issues until they were already in significant trouble.² Some of the consequences have been:

1. Significant unevenness within entries. There are different ideas about primary audience [discussed in the MDF manual in section §4.2]; different conventions for order of presenting information, formatting, orthography, glossing, definitions, naturalness and vividness of example sentences, cross-referencing strategies, etc. [MDF §2]
2. Doubling of effort. Entries for the same headword are made on multiple computers. This is further complicated by part of the entry on computer A being very well done, and a different part of the same entry on computer B also being very good. So a blind delete, or a blind merge doesn't solve the problem. The entries have to be edited by hand to preserve and merge the best of both into a single entry.
3. Unevenness across entries. Some entries and sense descriptions are at early draft stages and have been seen by only one person. Others are fully processed and have the input and consensus of the whole team that they are ready for publication.
4. Regression and mangling. A team member with limited skills and perspectives (accidentally) deletes or modifies information in a fully processed entry that they don't understand. There are no mechanisms for retrieving the earlier version, nor for preventing this kind of thing from happening. How many masterpiece entries are lost this way?
5. In trying to merge entries from separate files,
 - a) sometimes whole files are accidentally deleted,
 - b) some carefully worked through entries are overwritten by simpler entries,
 - c) some newer entries are overwritten by older ones,
 - d) some entries occur twice in the resulting file without being caught by an editorial process.

² One noticeable pattern that has emerged over and over is how many native speakers of minority vernaculars who are highly educated and reside outside their homeland have a desire to see dictionaries in their languages. But it is also apparent that many of these have no reliable sense of the grammar, vocabulary or usage. They are not reliable contributors to the dictionary. Just as having formal training in linguistics does not make one a lexicographer, so also just being a native speaker of a language does not make one a lexicographer. Both need training and mentoring in the specialized issues of lexicography to produce useful and reliable dictionaries.

6. Overwriting good files with bad ones. Because of not paying attention to filenames, version control, and back-ups it is quite common to overwrite newer fuller files with older less fully processed files.
7. Increased frustration and strained relationships.
8. Additional delays in getting the dictionary printed.

3. Team dynamics and roles

There are many skills required to compile a ‘good’ dictionary. When working as a team, these skills can be spread among several team members and do not all have to be encompassed in a single individual. With training, some team members may turn out to be exceptionally gifted at:

1. Having a rich vocabulary and being able to retrieve it. [various strategies in MDF §6.1]
2. Being able to distinguish homonyms, and different senses (polysemy). [MDF §6.3]
3. Coming up with good example sentences that are vivid and illustrate the different senses clearly. [MDF §6.2]
4. Phrasing things naturally. Surprisingly, many native speakers phrase things awkwardly, or force the grammar artificially when phrasing things for a book or on computer. It is also surprising how many native speakers are willing to put non-existent words or compounds in the dictionary (often artificially forcing the language to match the national or some other dominant language). Many non-native scholars have the same problem, and also often phrase things ungrammatically. A dictionary team needs at least one native speaker with a good sense of naturalness and grammaticality. As their intuitions are repeatedly confirmed by other speakers in the community, the whole team can grow in their confidence in the reliability of the intuitions of that individual.³
5. Structuring information in an entry and in the database. [MDF §5]
6. Editorial/copy-editing skills for consistent interaction with the dictionary program, and satisfying results. Some people are gifted in this area and others are hopeless.
7. Basic competency in using computers, file management, TOOLBOX, and MDF.
8. Handling dialect information and special registers. [MDF §6.5, §8.4]
9. Having a sense of current usage, as well as traditional, ritual, or archaic usage.
10. Cross referencing with lexical functions. [MDF §7]
11. Parsing the morphology (where relevant).
12. Part of speech issues. [MDF §9]
13. Reversals. The lexical database can produce an automated and detailed reversed finderlist if the team understands the process and possibilities. [MDF §2.3]
14. Labeling semantic domains for alternate retrieval. [MDF §6.4, §10.1]
15. Developing (gathering, recording, and keyboarding) a large corpus of natural texts for concordance searches in TOOLBOX for:

³ Joe Grimes (pers. comm.) observes, “This growth can be enhanced by having the dictionary accessible on the Web to qualified individuals not on the team, with a mechanism like Wikipedia for getting feedback without changing the master files.”

- a) example sentences,
- b) sense discrimination,
- c) analyzing the semantics of grammatical functors. [MDF §4.4]

In addition, many dictionary teams need access to experts to help them with:

- 16. Etymologies. [MDF §8.3]
- 17. Borrowed words (loans). [MDF §8.3]
- 18. Flora and fauna. [MDF §8.1]
- 19. Kinship systems. [MDF §8.1]
- 20. Material culture. [MDF §8.1]
- 21. Folk taxonomies. [MDF §8.1]
- 22. Cosmology and worldview, etc.

These experts may be part of the team already, or may be consultants whose regional and technical expertise is tapped into. Most people on the team would be expected to develop competency in several areas and be exceptionally gifted in at least one.

In addition, at least one of the key decision-makers needs to have or develop understanding and competency across a broad range of the skills listed. It is not recommended to have a ‘figurehead’ at the helm of the team who is there simply by virtue of position in society but does not understand the technical and audience issues. I have seen some silly decisions made out of good will on the part of others on the team, and ignorance on the part of a figurehead leader who is willing to make snap policy decisions without becoming informed on the issues.

4. Advantages of using several native speakers

We are all products of our histories. Therefore we have greater and lesser exposure to different areas of life. One may know a lot about farming, but very little about fishing. One may have a vast knowledge and awareness of several dialects and registers, while another may know only the dialect they grew up with, and only the common register. One may be gifted at getting a basic entry laid out on the computer or on paper, but be hopeless in wording example sentences naturally. Everybody has limitations. So where possible, it is advisable to have several native speakers involved on the team—provided they can all be adequately trained.

5. Some solutions

There are a number of solutions that can help reduce some of the problems mentioned above. Each also has their complexities and limitations.

5.1 Solution #1: only one at a time

The simplest solution for regaining control of the chaos is to go back to a system where everything is bottlenecked through one person on one computer.

ADVANTAGES: This ensures that everything funnels through the person who ‘knows’ the lexicography issues, the language, the culture, the format, the computer program, etc.

DISADVANTAGES:

- 1. The rest of the team has reduced involvement and hence reduced ownership. The dependency structure also means that other team members may never grow into mature

lexicographers who can later take the dictionary to the next level, or train others in neighboring languages.

2. The dictionary cannot progress if key members are scattered in different locations.
3. Work on the dictionary is effectively limited by the abilities and limitations of the person in control.
4. Work on the dictionary is limited to the combined schedules and interruptions of key team members being able to get together at the same time.
5. There is a good chance that significant material and entries not known or accessible by the person in control will not make it into the dictionary.
6. Incapacitation or death of the key person may be the death of the project as well, because nobody else knows how to carry it forward.

5.2 Solution #2: version control and file-sharing software

There are varying degrees of sophistication possible here. More options are becoming available all the time, so the key here is access to a computer geek who is aware of what is out there, how to get it, and how to use it. Most of these also assume reliable high-speed internet connection, which is not available in many situations. Some options have been tried with varying success, and with varying degrees of satisfaction or disappointment.

1. Use NET MEETING (or its functional equivalent). One computer has the master copy of the lexical database. At the end of a session the newly modified master is redistributed to the others for safe-keeping.
2. Use Version Control Software, or a Document Control System.⁴ These often work as a master document with a ‘check in’ – ‘check out’ system. For many systems, only one person can work on a file at a time. Some of the software is fairly expensive for a dictionary team operating on a shoestring.
3. I know one team working at different locations on different computers. They have paid for a Web service to discuss things on-line in real-time by enabling the remote computer to view the screen on the master computer, and allowing the person on the master computer to turn over control of the keyboard to the remote computer to type in suggestions for on-going discussion and evaluation before they lock it in. They are using this approach successfully, and point out that the cost of the software necessary to do this on a regular basis is significantly cheaper than the travel costs of getting together. They also note the time formerly wasted for travel is now productively used on the dictionary. And it gives them added flexibility for scheduling. This approach assumes a degree of computer skills or computer support that may not be available to some teams.⁵

⁴ Some that are available at the time of writing are: SVN (also known as SUBVERSION), CVS, PERFORCE, and SUPERVERSION. Joe Grimes describes the system they are using as follows: “Only one person, as Editor, has Commit privileges. So people working on a text, Update from the repository before they go to work. Then they email what they did to the Editor, who reviews it for coherency with the grand plan, edits if necessary, then Commits it to the repository. We plan to add editors with Commit privileges as some members of the team learn the editorial ropes. It requires high speed Internet, but could be done on dialup if we used a smaller file size.”

⁵ A couple of Web services available for desktop sharing are LogMeIn.com and GatherWorks.com. Some allow for four or five people to view the computer that paid for the service, at no cost to them. For voice dialogue, many teams are also using Skype.com as their voice over IP channel, free with other Skype users anywhere in the world, and capable of handling up to 5 users at once. Skype works best with a high-speed connection.

While these approaches address a number of file management and version issues, they still do not address all the issues listed above in section 2. They are certainly partial solutions for some teams.

5.3 Solution #3: master file and satellite files

The procedure recommended here is not dependent on additional software or reliable Internet connections, but on defining and managing a process, an on-going communication. As such, it has been shown to work successfully in a variety of situations and allows for a large degree of flexibility. It empowers the team to work in the areas of their strengths. It is how I work.

Briefly, there is a MASTER (MATRIX, MOTHER, MAIN) FILE in which all the entries have been fully processed and is incremented through additional versions as other fully processed entries are added in. This requires a FILE NAMING SYSTEM to keep the various stages distinct. Everybody has a copy of the current MASTER FILE (computer version or hard copy), but nobody is allowed to add or change material directly on this file on their own.

There are separate SATELLITE (INPUT, DAUGHTER, FEEDER) FILES for adding new material, correcting existing entries, or expanding existing entries. Each team member has their own satellite file that is incremented through different versions. This requires a robust FILE NAMING SYSTEM to keep the various team members and stages distinct. For efficiency it is useful to divide tasks or domains to minimize overlap. For example, one person works on fish, another on agricultural plants, another on kin terms. Or one works on cutting verbs, another on carrying verbs and related implements. Each person can work as much as they want, and flesh out the entries to the best of their individual abilities. Some team members may not put in much more than simple glosses. Some may have very sophisticated entries with complex internal structure. Part of this depends on (on-going) training and experience.

It has been observed that the most difficult thing about computing is not data entry, but file management. It should not be surprising that some team members may find file management issues difficult or intimidating, or accidentally delete critical files. So it may be wise to designate someone who has access to many of the team computers to assist with making back-up copies of files, renaming files to the next increment stage, and emailing off copies of files at given stages for further processing.

On a periodic basis, all members of the team (or else just key members with critical skills) need to have blocks of time to work together to PROCESS THE INFORMATION IN THE SATELLITE FILES. During this time they check, refine, and add to the information structure of entries, the glosses, sense discrimination, confirm or refine the naturalness of example sentences, add etymologies, add lexical functions [MDF §7], tighten up the national language renderings, tighten up the English (or other glossing language) renderings, etc.⁶

The team can choose to process all the entries in a whole satellite file before MERGING IT INTO THE MASTER FILE. Other teams will prefer to copy each entry from satellite files into the master file as that entry becomes fully processed and the team present agrees that it is ready, because some entries will need to be set aside for further processing before they can be merged in. (In TOOLBOX this can be done easily with the DATABASE, COPY RECORD feature.)⁷

⁶ One team has formulated a general rule as follows: “Anybody on the team can suggest an improvement to anybody else’s work. Anybody who proposes an improvement should have a reason for it.”

⁷ If things get confused because of the same filename on different computers, TOOLBOX (and other programs) has a FILE COMPARE utility which can highlight the different text strings between similar files. Some programs can track who made what changes in which fields in an entry.

After these satellite files are processed and merged into the master file, the new master file is given a new filename, REDISTRIBUTED to all team members, installed on all the relevant computers, and the JUMP PATH is adjusted to look in the new master file, not the old one. (Again, this may be a computer skill that is beyond some team members. They may need help to do this, or need someone to do it for them.)⁸

A NOTE on working with SHOEBOX/TOOLBOX:⁹ These programs can be set up for each team member to have the current master file and their own satellite file viewed side-by-side. They can be linked through the JUMP PATH so that any attempt to add a new entry can automatically search through both files to see whether it is already in one of them. This is really helpful once there are more than a few hundred entries in the lexical database. It also helps for discovering and indicating homonyms. Each team member can also have the entire TEXT CORPUS on their computer for interactive CONCORDANCE searches through TOOLBOX. Corpus based lexicography gets us away from building dictionaries around word lists, whose semantics is usually that of the glossing language, not any of the languages being studied. [MDF §4.4].

The process described above can be visualised in more detail as follows:

If several people have been making entries on different computers or different files, before getting things organized, back up all files for safe-keeping in case something gets inadvertently deleted or mangled in the learning process.¹⁰ Open new empty satellite files for each team member to work on from this point forward. They should not contain the material already in their previous file.

Establish a consistent file naming convention. One that seems to be quite flexible is:

Language Abbrev-LEX-Person-Version.db
 Iranun-LEX-Husein-A.db
 Kupang-LEX-June-G.db
 Kupang-LEX-Chuck-D.db
 Ama-LEX-Roni-B.db
 Buru-LEX-Chuck-J.db
 Bai-LEX-Theo-A.db

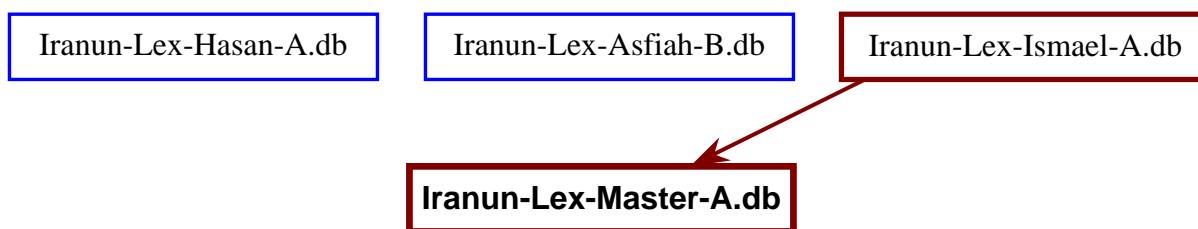
Examine all the files, choose one to become the initial master file. It may be the one with the most entries, the one with the most sophisticated information structure in entries. or the one with the best understanding of how the lexical database interacts with the computer program for getting a formatted printout and reversal. Reassure everyone that all the information in their files will get into the master file, so nobody will feel slighted.

⁸ One way of distributing a new copy of the master file is to place it in a Web repository where team members can download it. For work-in-progress it may be wise to put this on a secure passworded site. For certain computer programs the appropriate settings files could also be placed in the Web repository so that the new versions of the file come up automatically. That reduces the setup on the satellite computers to just copying the right files into the right folders. Version control software avoids having to keep renaming the master file. The repository handles all that automatically. It also has a mechanism by which you can view any earlier stage, so if something gets wiped out you can copy it from the earlier stage to the current stage.

⁹ TOOLBOX is basically a later version of SHOEBOX with enhanced features. Among other things it is Unicode compatible. It can be downloaded from <http://www.sil.org/linguistics/computing.html>.

¹⁰ It is also useful to save a copy at one or more other locations (in another country) in case of fire, theft, lightning strikes, leaking roofs, riots, cockroaches, rats, dust, or mold. A Web repository is another way to achieve the same goal.

Figure 1: Choose one file to become the Master file

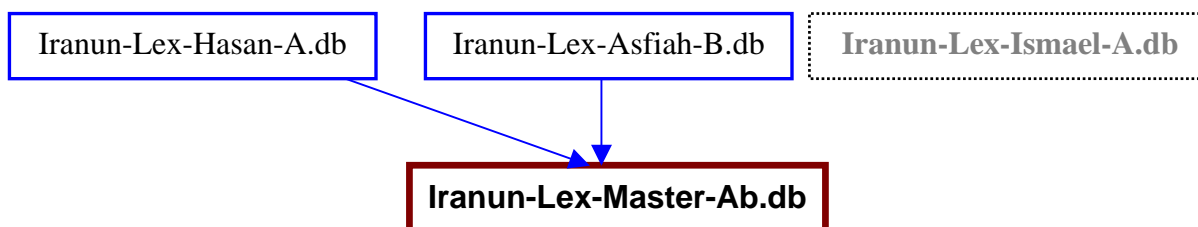


The file chosen to become the master file will be examined entry-by-entry together as a team, refined and expanded together. This is a team learning exercise. It is wise at this stage to begin writing down decisions made as tentative guidelines that the team wants to follow. This will reduce the amount of “everyone does what is right in their own eyes” which results in uneven chaos. This way people who join the team at a later stage can also be brought up to speed.

(**) The other files can then be compared with the new master file and refined as described above. During this process the two files can be compared side-by-side in TOOLBOX (Use WINDOW, TILE SIDE BY SIDE). The refined and expanded entries from these other files can be copied or merged into the master file. (Using DATABASE, COPY RECORD...)

If the same headword appears in both files, check carefully to ensure that all the information has been added to the entry in the master file.

Figure 2: Copy fully processed entries from other files into the master file



When information from all the files has been examined, refined, expanded, and merged, then the new version of the master file can be placed back on each of the computers with a new filename. It needs to be clear that this new master file is fixed, and may not be edited or expanded by any single individual.¹¹ All editing is done on separate satellite files. Any suggestions for changing or refining information on the master file must wait for the next round of group processing.

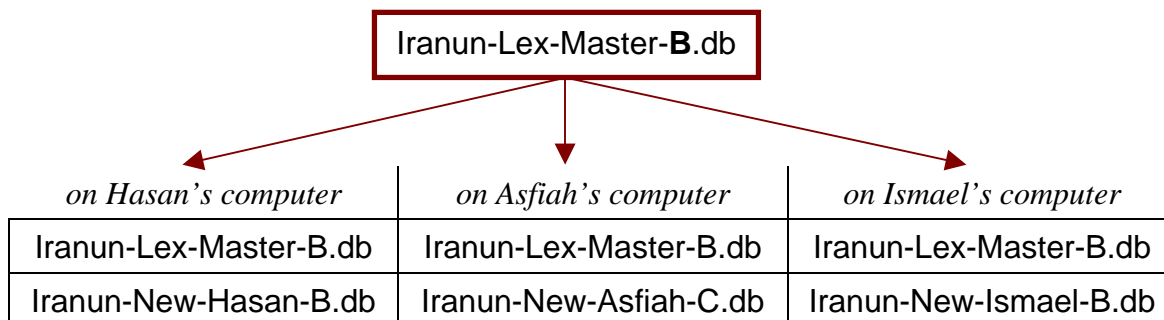
The process for expanding or editing entries that already exist on the master file is to copy that record from the master file to the satellite file (use DATABASE, COPY RECORD...), and then edit it on the satellite file. When it comes time to process that entry for merging back into the master file, the suggestions can be discussed and the old entry can be deleted from the master file and replaced by the new one at that stage.

Once all the entries from all the files in that round are fully processed and merged into the master files, new empty satellite files are set up on each computer for the next round of work. Each file is given a different filename (as per the suggestions above). The different areas for each person to focus on needs to be discussed and communicated to everyone on the team to minimize

¹¹ The Master File can be made to have ‘Read-only’ file attributes, but that adds other complexities. In a Web repository, granting Commit access only to qualified individuals takes care of this problem. Everybody else makes proposals for change, and sends them to one of the ones who really knows the ropes.

duplication of effort. Agree on tasks, targets, and when to get together for the next round. Aim for 200 new entries each? 500? Get together routinely every six months for 2 weeks of intensive work?

Figure 3: Copy new expanded master file onto each computer and open new empty satellite files



When the appropriate files have been copied to each computer, TOOLBOX can be set up to look first in the master file, and if the headword is not found there, it can then look in the new satellite file. This can be done through: DATABASE, PROPERTIES, JUMP PATH, DATABASES IN PATH, enter the filename of the master file first (e.g. Iranun-Lex-Master-B.db) and then the filename of the appropriate satellite file second (e.g. Iranun-New-Asfiah-C.db).

At the next round, the cycle of refining and expanding the entries in each of the satellite files and merging them into the master file can be done again (begin from the double asterisk **).

Don't forget to back up all computer files BEFORE beginning to process them at each round.

Don't forget to save copies at a different location 'off-shore'.

Don't forget to give all files a new filename at the end of each round, and then not do any editing in the old files. The safest way to ensure this is to delete (after saving archive copies) or move the old files to another folder on all the computers.

The process suggested here sounds more complex than it is.

6. Other issues when working as a team

One can define processes, file naming conventions, and have the best lexicographers and computer programs in the world, but if the interpersonal relationships and cross-cultural communication skills are not there, the dictionary will probably not be seen through to publication. While it may seem obvious, enough dictionary teams have gotten in trouble that it is worth saying.

Interpersonal and other skills that work well on a dictionary team include: the ability to listen to others, the ability to genuinely value what they say, a willingness to give in to others after everything is fully discussed, the ability to do careful detailed work, and a clear understanding of the primary audience for the dictionary.

In addition to the skill sets listed in section 3 above, several other skills are valuable to have on the team. These include:

- management/coordination: not for telling people what to do, but for arranging times to get together, locations, facilities, meals & snacks, paper, printer, keeping track of finances, etc.
- someone who is gifted at seeking funding for team activities and necessary equipment.

- someone who is gifted at marketing and promotion.
- Some people start working on a dictionary without anyone having figured out the best way to write the language. So under these circumstances at the early stages it can be helpful to have access to someone who has training, expertise, and experience in working with communities to help them work out and test a practical orthography that they can use in a sustainable and productive way. Some basic principles can be found in Smalley (1963). Examples of how these principles are applied to a number of Austronesian languages of eastern Indonesia can be found in Grimes (1999).
- Every language has its challenges. There will be some headwords that are so complex they defy any obvious organization of information structure. It is helpful for the team to have access to an expert lexicographer who can advise them on some of these trickier entries, and help clarify things along the way.

Awareness of the issues discussed in this paper and following the suggestions (with modifications appropriate to the situation and the capabilities of the team) will go a long way toward reducing chaos and stress on the team. Many teams wanting to do dictionaries forget that becoming a team of skilled lexicographers is a PROCESS that requires on-going training, continuing discussion, and a bit of trial-and-error. It is easy to do a dictionary poorly. It is quite a complex task to do a dictionary well that will be a service to both the local community and to others for decades to come.

References

- Aplugi, Lazarus, Charles E. Grimes, Ayub Ranoh, and Michael Sina, compilers. 2000, and in process. *Kamus Pengantar Lii Dhao (Bahasa Ndao)*. Kupang: Artha Wacana Press.
- Bani, Heronimus, Charles E. Grimes, Welfrid Fini Ruku, compilers. in process. *Kamus Pengantar Bahasa Amarasi*. Kupang: Artha Wacana Press.
- Beeh, Mes, Stuart Cameron, Charles E., Grimes, and Pieter Pong, compilers. 2001, *Kamus Pengantar Bahasa Helong*. Second edition. Kupang: Artha Wacana Press.
- Coward, David F. and Charles E. Grimes. 1995, 2000. *Making dictionaries: a guide to lexicography and the Multi-Dictionary Formatter*. Waxhaw: Summer Institute of Linguistics.
- [Since 1997, Windows and Mac versions of MDF has been upgraded and incorporated into the Shoebox program (versions 3.07 and above). Full electronic version (PDF) of Coward & Grimes 1995 released with Shoebox version 5.0 in September 2000. See also <http://www.sil.org/computing/shoebox>. MDF manual in PDF also available on CD-ROM on *LinguaLinks Library* v2.0. See http://www.ethnologue.com/LL_docs/show_contents.asp?bookshelf=Linguistics]
- Grimes, Charles E. 1999. Implikasi penelitian fonologis untuk cara menulis bahasa-bahasa daerah di Kawasan Timur Indonesia. In Soenjono Dardjowidjojo and Yassir Nasanius, eds. *PELBBA 12: Pertemuan Linguistik (Pusat Kajian) Bahasa dan Budaya Atma Jaya Kedua Belas*. Kanisius: Yogyakarta, Indonesia. pp. 173–197. [“Implications from phonological research for ways of writing vernacular languages in eastern Indonesia.”]
- Grimes, Barbara Dix, Charles E. Grimes, Yohanes Hukunala, Wesly Lesnussa, and Otniel Seleky, compilers, in process. *An encyclopedic dictionary of the Buru language*.
- Grimes, Charles E., Bernadus Lado, Thomas Ly, and Simon Tari, compilers. 2003, and in process. *Kamus Pengantar Lii Hawu (Bahasa Sabu)*. Kupang: Artha Wacana Press.
- Jacob, June, and Charles E. Grimes, compilers. 2003. *Kamus Pengantar Bahasa Kupang*. Second edition. Kupang: Artha Wacana Press.
- Smalley, William A. 1963. How shall I write this language? In William A. Smalley, ed. *Orthography studies*. London: United Bible societies. pp. 31–52.

The preceding document was presented at the Tenth International Conference on Austronesian Linguistics (10-ICAL). To properly reference this work, please use the following format:

<LastName>, <FirstName>. 2006. <PaperTitle>. Paper presented at Tenth International Conference on Austronesian Linguistics. 17-20 January 2006. Puerto Princesa City, Palawan, Philippines.
<http://www.sil.org/asia/philippines/ical/papers.html>

For other papers that were presented at 10-ICAL, please visit <http://www.sil.org/asia/philippines/ical/papers.html>.