

# Introduction to linux LCC tutorial (Handout for first part of Tutorial)

David Dowty (this handout), Adriane Boyd

## 1 INTRODUCTION

### 1.1 WHAT IS LINUX?

a variety of unix systems that runs on PCs (and now on most other hardware too) and is open-source (free) software. More accurately called *GNU/linux*. (cf. *kernel* versus other system software and applications). GNU (“Gnu’s Not Linux”) Project, sponsored by Free Software Foundation is the producer of the most important system software on a linux system besides the kernel (purely linux); X Window system

linux	<a href="http://www.linux.org">http://www.linux.org</a> <a href="http://en.wikipedia.org/wiki/Linux">http://en.wikipedia.org/wiki/Linux</a> , <a href="http://en.wikipedia.org/wiki/Linux_kernel">http://en.wikipedia.org/wiki/Linux_kernel</a>
GNU Project:	<a href="http://gnu.org">http://gnu.org</a>
X Window System	<a href="http://www.x.org">http://www.x.org</a> , <a href="http://www.rahul.net/kenton/xsites.html">http://www.rahul.net/kenton/xsites.html</a>

### 1.2 WHAT IS ITS HISTORY?

(From Linus Torvalds 1991, to today’s widely distributed and complex development teams, involving many many coordinated volunteers.

What are *distributions*? There are many different distributions of linux (Ubuntu and other Debian-based distributions; RedHat, Suse, etc.); cf. also BSD unix.

Ubuntu distribution:	<a href="http://ubuntulinux.org">http://ubuntulinux.org</a>
Debian distribution:	<a href="http://debian.org">http://debian.org</a>
linux history:	<a href="http://www.linux.org/info">http://www.linux.org/info</a>

Who uses it today? (individuals with various reasons to be interested; scientists; web servers; file servers (e.g. OSU ling); industry (e.g. CGI for films — *Titanic*); governments (Munich, Paris). Embedded linux systems (e.g. Tivo; see <http://linuxdevices.com> for more.)

There are now about a half-dozen magazines devoted to linux, including *Linux Journal*, *Linux Magazine*, hundreds of web sites.

(Search under linux on Yahoo for big list of sites, magazines, etc.)

### 1.3 WHAT IS OPEN SOURCE SOFTWARE?

(The GNU GENERAL PUBLIC LICENSE, a.k.a. GPL LICENSE or GPL) and why it is needed to protect open source software; Philosophy of open Source software.)

Gnu Public License: [http://en.wikipedia.org/wiki/GNU\\_General\\_Public\\_License](http://en.wikipedia.org/wiki/GNU_General_Public_License),  
<http://www.gnu/licenses/>

## 2 PROS (VIS-A-VIS MS WINDOWS)

Mac OS X (for Macs) and FreeBSD (for PCs), maybe others, are also alternatives to consider; some of these points are also advantages over those, most are not (e.g. cost).

### 2.1 ADVANTAGES OF LINUX FOR OSU LINGUISTICS STUDENTS SPECIFICALLY

- Same system (essentially) as unix system on the ling network (term ”\*NIX SYSTEMS” includes unix (Sun’s Solaris, etc.), linux, FreeBSD (<http://www.freebsd.org/projects>), others.
  - Can help you learn unix more quickly and easily. (Not only unix itself but also applications like emacs, latex, etc. will work exactly the same as the unix version on ling.)
  - Interfacing with ling system easier (e.g. remote logins (both directions), ssh, assuming you have wide-band internet access at home: remote X, file non-conversion).

- Some linguistics applications available in unix/linux only, and/or the linux versions is better than the MS Windows version. Examples from my areas include GRAIL (categorical Grammar grammar implementer), JAPE (framework for building theorem-provers or tutorials for any deductive sytem you like, or can create). In comp ling, detmar mentions CONTROLL/XTROLL and TRALE/GRISU, two leading HPSG implementation platforms (URLs for these available upon request.) (⇒ Adriane will describe some others later)

Also: more software customizations and extensions specific to lingusitics exist (in part because of greater configurability). LaTeX is a good example (w/ caveat).

- Helpful if you’re specializing in comp ling and would find it useful to know more about compiling software and about system administration of linux/unix.
- (A Non-disadvantage): many people around here use it, easy to get help, share techniques and tips, customizations, etc. A number of important applications (notably latex incl. xdvi, gv, etc.), can in fact be used in MS Windows versions but usually aren’t by linguists, for various reasons.
- There is a broad (on-line) community of people working with \*nix systems which (as Detmar reminds me) has a very different spirit of cooperation from what you seem to find with Windows; this has concrete and practical consequences. Within linguists specifically, you likewise find this community spirit for developing and supporting \*nix software for linguistics.

Caveat: these advantages apply most of all to students working in S-track areas (syntax, semantics, computational linguistics, pragmatics) somewhat less so to phonetics and psycho-ling., probably least in Historical, Socio- (?).

## 2.2 GENERAL ADVANTAGES FOR ALL/MANY KINDS OF USERS (NOT SPECIFIC TO LINGUISTICS)

- Linux system software is free.
- Linux applications are free (but you might want to pay very rarely for commercial extra software)
- Configurability: apparent user philosophy of Windows:

“The user is assumed to be computer-naive: she/he can’t mess with software without mucking it up (so the system software must be protected from him/her), she/he will never want to know more than the minimum information about how the software really works, and will accept a one-size-fits-all user interface”.

Vs. \*nix philosophy:

“The user IS able to understand software (or will be when given enough information), wants to have control over it and its user interfaces, to always have the choice of changing most any aspect of it.”

- A greater variety of options can usually be re-configured in linux applications than the corresponding Windows application.
  - For almost every category of utility and application program (starting with the desktop itself), you have multiple alternative programs to choose from. Examples: KDE vs. Gnome vs. XFCE vs. other desktops/window managers. (A historical reason: linux developers often write new versions to improve on them, and for the fun and the challenge of it.)
  - HOW to do all the possible configuration changes is easier to find out with linux (Man pages, HOW-TO files, other documentation in system as well as much more on line — in addition to those familiar pull-down Help Files.)
  - Possibility of re-compiling software from source code, after making your own possibly major changes to it. ( ⇒ Adriane)
- The Internet and Internet Protocols were originally developed to run unix systems, so network configuration and use is (I think, at least) somewhat more straightforward and flexible in linux than Windows.
  - Security: (a *big* issue!)
    - a historical difference: unix originated as a multi-user, networked system, so protecting users’ data from other users and protecting system software from users’ mistakes always had a high priority. (Note: ”root” (administrator) vs. ”ordinary user”). Windows originated as a single-user, non-networked system, so networking and network security features were added on later – much later)
    - \*nix has a hierarchy of access to files (and programs and hardware devices):
      - \* users (root, ordinary users, programs); groups.

- \* file permissions/protections (read, write, execute – specified for user (owner) vs. group vs. others).

- \* a hierarchy of programs and what resources they have access to (privileges). (one reason why email viruses don't work).

Principle: users and software are given access only to the files and resources they specifically need, never any more. (An intruder or trojan that gains control over just one program can do only limited damage, similarly for just one user's account—which is why root access is stringently protected.)

- Windows has high risk from email viruses, downloaded software, Internet Explorer.
- Security updates to linux software are in general made more quickly, often much more quickly (contra rumors spread by you-know-who).

[Is there less risk from browser spyware and adware with linux? Possibly, not clear how much, to me anyway. Spam annoyance depends on so much on what filters you choose to use that no generalization may be possible about systems..]

- A DUAL BOOT SYSTEM is possible (you can set up your computer to operate on both Windows and linux)—another very important thing to realize! Best of both worlds. PARTITIONS and of FILE SYSTEMS TYPES. Files in Windows partition are readable by linux system, not vice-versa.
- Certain linux applications don't exist for Windows, or are superior in linux versions. (Scientific software in general is more often found in unix version; historical reasons for this — use of mainframes arose before PCs or Windows existed). Number of applications ported to linux is growing rapidly, including so-called "clones" of popular commercial software. Example: Microsoft Office → OpenOffice. PhotoShop → Gimp, IM →. Sometimes linux version of a commercial program is free while WIndows version of it costs money.
- Overall, linux software is (widely believed to be, by people who should know):
  - more stable (our historical reason: with multi-user system, crashes and down time must be kept to a minimum; it should be possible to make changes to system software and hardware without rebooting the system or interrupting other services).
  - It is often held that linux software is less buggy (at a high level), more streamlined and less bloated (one historical reason: the first linux developers had older slower computers, but what's important today is that quality of software is naturally a priority for linux developers, lack of deadlines). (Fallacious argument: "Windows was developed by full-time professionals, linux by amateurs in their spare time, so Windows must be better".)
  - Updates, both to system and applications, are made available more frequently (both security updates and other bug-fix updates), contrary to certain rumors.
- (A non-disadvantage): In current linux distributions, installing new software and making it work is "painless"; installation of system itself is fairly "painless", so is configuration of new hardware (mostly); updating software is automatic.

(Two real advantages): (a) Most or virtually all the software applications you use will be available from a single source, the Debian (or debian derivative) software archive (in latest version), you don't have to locate each individually. (b) In a Debian-package-based package

system (which includes Ubuntu), you can update ALL the software on your computer to its most recent and most secure version with a single command (!)

- Documentation (for everything) is available on-line and free, easy to locate, usually well-written and detailed; much newsgroup discussion is accessible (by developers as well as users).
- (A non-disadvantage:) One fix for a number of instances of the need for software that only exists in Windows version is a Windows emulator—utility that creates a virtual Windows environment running inside the linux operating system.

WINE (sometimes written ‘Whine’) = ‘Wine Is Not an Emulator’, a.k.a. ‘WINDows Emulator’) is the standard (free) linux emulator (though doesn’t really work with lots of Windows software).

*Crossover*, made by Codeweavers, works quite well for many more Windows applications, also runs some browser plugins that don’t have linux versions, though is commercial, proprietary software: costs \$39.95 in limited-feature 6-mo. version, \$69.95 in complete, 12-mo. version.

Crossover: <http://www.codeweavers.com>

- Philosophical/Ethical: Resisting the “Evil Empire”; belief in the importance of open source software in general.
- There are many basic features of the \*nix environment (which can all be used in a terminal window) that have significant advantages; including powerful search tools such as grep (which let you search big files and directories using regular expressions, counts of some string in a large file), and piping of commands (feeding output of one command directly into another command, etc.); I really like the terminal command-line editing features that are as flexible as those of a word processor. BUT since there are separate tutorials on Unix, we won’t talk about these features here.

### 3 CONS: (AS COMPARED TO STICKING WITH WHAT YOU ALREADY HAVE NOW)

- Investment of the time to do initial installation and configuration (not a lot with today’s linux systems, was formerly significant).
- Investment of time in learning to use linux (more considerable amount): Learning curve can be somewhat steep (depending on how deep you want to go), though you can get by pretty well initially with just the familiar Windows-like desktop features of Gnome or KDE (etc.) and their own interfaces to other applications, if you want to.

(But, people who are good at linguistics are almost always quicker at adapting to novel computer systems than the average user, even linguists with little prior computational training.)

- Some amount of system administration knowledge and attention to system maintenance will always continue to be needed, on a more or less regular basis, to keep your linux system in the most effective working order.

- Despite the appearance of more and more Windows clones, etc., there is still lots Windows software that does not exist in \*nix versions—this happens more rarely with software used by academics/scientists than with recreational software. (And as far as I know, there is little or no important linguistics-specific software that runs on as far as I know).

The Windows emulation programs mentioned above still won't run every Windows application.

And emulators are not really suitable for running large, interface-intensive Windows applications for a long period of time.

Rebooting into Windows is always an option (but becomes time-consuming if you have to go back and forth repeatedly).

- Sometimes the linux version of an application has fewer features than the Windows version (example: Acrobat Reader).
- A few browser plugins don't exist in linux versions. Also, a few web sites will only work if you are using Internet Explorer (for no particularly good reason)—a problem if you use firefox, opera, etc., on any kind of platform.
- The one real security threat you do have to deal with in linux, if you want to allow contact to your system to be initiated from the internet that is, is risk of intrusions (break-ins by crackers)—though good firewalls and other protection exist. (Contact from inside to outside is not a problem.)
- “We don't support linux” — is something you often hear when you call tech support for ISPs, hardware, web site issues, etc. (Usually there is really nothing about your problem that is really due to linux *pe se*, but there is a communication problem to be dealt with somehow.)
- Some people find that a linux system turns out to be a lot of fun to play around with, so it could take time away from other activities (important to identify such activity as “recreation time”, not time spent working on linguistics itself—but the boundary is often hard to draw).

Finally, the option of the Mac OS X (Tiger) operating system and environment should be also be considered if (i) you already own a Mac, or (ii) you are due for a replacement for your present computer, can afford a (fast) Mac, and are happy with Mac software as your alternative to linux (rather than Windows as the alternative). The Mac OS X is a system like traditional Mac in many ways, but is actually a unix system at the core, will run most any \*nix application, especially scientific ones, and even an X Window interface instead of Mac's usual one. (Ask Detmar for information and his evaluation.) It costs about \$125 at present.

## 4 LIVE DEMOS

I will demo the basic Gnome linux desktop, network interface and Windows interface, using fang.ling.ohio-state.edu, the computer that is in my office and part of the ling network; it runs Debian linux (3.0 currently). Then Adriane will discuss some of the important major linux (\*unix) applications and will use her lap top with a “live CD” form of Ubuntu linux.

1. login screen (gdm), note choice of session manager (desktop).

2. main gnome desktop: point out panel, note you can add/delete/move additional panels, add/delete icons in panel (uncover desktop icons). Notice that this looks and acts very much like a Windows desktop (but it can be reconfigured more extensively).

(If time: show screen shot of the default Debian desktop you get as a new user.)

3. Note Terminal windows (X-Term)—which you will probably use more frequently than the DOS-type console in a Windows system.

2. (Probably not shown: linux virtual terminals; screen resolution change via Ctrl-Alt +)

3. Multiple desktops (a very notable and desirable feature for most people!), show with both (a) selector in panel and (b) with “edge flipping” with cursor, and sliding window to adjacent desktop with cursor. (Might skip) Add a new desktop, rearrange them.

4. (Probably skip: Example of a GUI system Configuration gnome utility. Also, note root login.  
)

5. changing keyboard mapping.

6. Nautilus file manager.

7. (May skip: Using emacs directory as a file manager instead.)

8. Show that you can view the Windows partition directories and the files in them from linux system.

9. Demo WINE, with Tarski’s World as an example.

10. (May skip: install a new software package with synaptic (graphic interface to apt-get, the Debian package manager.)

11. Using ssh (Secure Shell), Log in to my linux computer at my home, over the internet, then run something in a term window (ls -l), then start an X application or two (maybe Emacs, xv), that is, the application will be running on my home computer but its window will be displayed here. (Using graphical interfaces remotely may or may not be a practical way to work at home, depending on your internet speed; VNC (Virtual Network Computing) may be faster than remote X-Windows; running applications remotely in terminal window is always fast.)

(WARNING: I demo this mainly to illustrate how you would log in to the ling system from your home computer—that is, in the “opposite” direction from what I’m showing, but works the same way. You can (but need not) also configure your home computer to allow logins to it from elsewhere on the internet (as I do here) and leave it on, BUT you should take a number of security precautions if you do this which are very important!)