

## The XTrans Tool

XTrans is a next generation multi-platform, multilingual, multi-channel tool developed by Linguistic Data Consortium (LDC) to support manual transcription and annotation of audio recordings. As human language technology expands its horizons, data providers must respond with infrastructure that supports rapid development of high quality, high-volume linguistic resources in a nearly unlimited number of languages and genres. XTrans provides new and efficient solutions to common transcription challenges and addresses critical gaps in existing tools.

### XTrans is...User-friendly

- intuitive, efficient GUI
- rich integrated QC features
- keyboard control for all tasks
- easy handling of multi-speaker overlap
- user-configurable keybindings and shortcuts
- language-independent text input and display

### ...Adaptable and Flexible

- unlimited audio channels per session
- most common audio formats supported
- interoperable with common transcript formats
- simple, flexible output format

### ...Developer-friendly

- multi-platform
- open source and common scripting language
- freely distributed

Designed with input from experienced human transcribers working with real world data, XTrans provides a flexible and intuitive graphical user interface for a multitude of speech annotation tasks including (virtual) segmentation of audio into smaller units such as turns and sentences; speaker identification; orthographic transcription in any language; and labeling of structural elements of the transcript including topics.

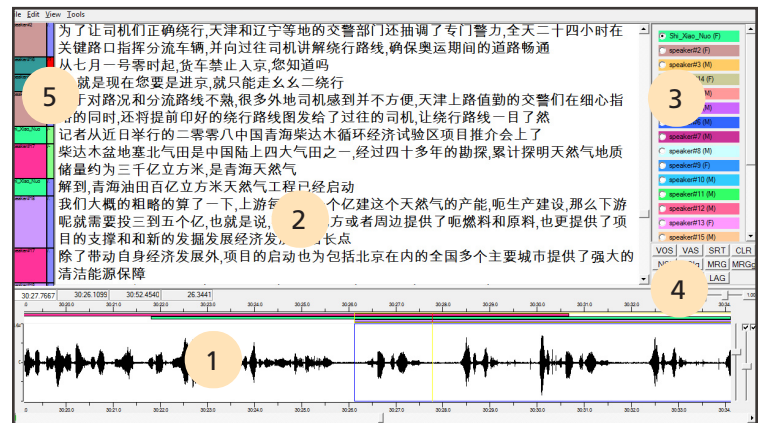
XTrans features embedded quality control functions, color-coded speaker and segment labels, user-configurable keybindings, and a simple output format that can be easily adapted to downstream processes.

## XTrans Exposure

Since its creation, LDC and its government, academic and industrial partners have used XTrans to generate over 3500 hours of time-aligned verbatim transcripts in a variety of genres and languages. These include broadcast news, talk shows, sociolinguistic interviews, telephone speech, massively multichannel meetings and more, in languages including Chinese, Arabic, Spanish, Russian and English.

## Design

XTrans' physical layout includes separate panels for each major function:



1. The **audio panel** is tightly integrated with other elements: as users label a segment (utterance, turn, etc.) in the audio file, a color-coded rectangle marking the duration of that segment appears above the waveform.
2. The **transcript panel** supports orthographic transcription.
3. The **speaker panel** lists all speakers by name or unique identifier.
4. The **widget panel** incorporates speaker identification and QC functions.
5. The **segmentation panel** displays segments of different granularity, like sentence units, speaker turns, or topic boundaries.



A screenshot of Arabic transcription involving multiple channels, overlapping speakers and bidirectional text

## XTrans Features

With an [intuitive interface](#), [user configurability](#) and [embedded QC functions](#), XTrans is optimized for high-quality, high-volume transcription tasks involving real world data.

Real data means messy data. Transcribing [multiple overlapping speakers](#) on a single channel can be a challenge for even the most experienced transcribers. XTrans eliminates cumbersome workarounds for this common phenomenon with Virtual Speaker Channels, enabling a virtually unlimited number of distinct speakers or other sound sources to be associated with the same audio channel.

Real data means complex data. XTrans allows transcribers to open an effectively [unlimited number of audio files for simultaneous transcription](#). Transcribers can switch focus between one, two or multiple speakers as needed -- a great feature for transcribing meeting room recordings!

Real data means global data. XTrans provides strong multilingual support, with [bidirectional text input](#) for languages like Arabic, Farsi, Urdu and Hebrew. XTrans is Unicode compliant and provides support for Windows and Unix multilingual input methods.

## Efficiency and Quality

[Realtime transcription rates have improved dramatically](#) in LDC projects using XTrans, with rates for some tasks cut by as much as half. XTrans' combination of intuitive tool design, user-configurable conventions and a novel use of keybindings to support all tasks including audio segmentation provides the

[optimal balance of efficiency and quality](#). Experienced transcribers rarely even have to touch the mouse, which not only speeds up transcription but also reduces hand and wrist fatigue.

XTrans also brings [key quality control functions](#) directly into the interface, giving transcribers the power to improve the quality of their own work.

## Flexible Formats, Developer-Friendly

XTrans components are written in Python and C++ and use LDC's QWave waveform display module. Even with very large files or multiple recordings, XTrans provides fast display and playback capabilities and supports multiple audio formats including .sph, .wav, .aiff, .flac and .ogg. Transcripts are output in a Tab Delimited Format (TDF), which is easily converted to other common formats and is readily usable by downstream manual and automatic annotation tasks. XTrans directly imports files from some XML formats such as Transcriber (TRS) format.

XTrans was designed to be easily extensible to new tasks; LDC recently extended XTrans to build QCTrans, an interface for creating and validating sentence-aligned parallel text starting from transcripts of spoken language.

## Availability

Best of all, XTrans is free! XTrans for Linux and Windows is released under GPLv3.

Download your copy of XTrans today from [www ldc. upenn.edu/language-resources/tools/xtrans/downloads](http://www ldc. upenn.edu/language-resources/tools/xtrans/downloads)