

**The LENA™ Language Environment Analysis System:  
The Infoture Time Segment (ITS) File**

Sharmistha Sarkar Gray, Charles T. Baer, Dongxin Xu, & Umit Yapanel  
Infoture, Inc., Boulder, CO

ITR-04-1

October 2007

LENA™ Hardware Model: LR-0120    Software Version: V2.3.0

Copyright © 2007, Infoture, Inc. All Rights Reserved

## **The LENA™ System**

The LENA language environment analysis system is a language monitoring and feedback system designed to provide information about the language environment of infants and toddlers to parents, clinicians, and researchers. The LENA System includes the LENA digital language processor (DLP) that children ages 2 through 36 months wear in the pocket of custom-made clothing. It records everything the child says and hears over a continuous 16-hour day. The audio data is transferred to a computer and analyzed by the LENA language environment analysis software. Parents can access automatically generated feedback reports to view objective information about their child's language environment. The Adult Word Count (AWC) report provides estimates of the total number of adult words the child hears, and the Conversational Turns (CT) report provides estimates of the total number of conversational interactions the child engages in with an adult. These reports permit AWC and CT estimates to be viewed as hourly, daily, or monthly totals. Daily AWC and CT percentile ranking estimates based on a normative database are reported in the LENA software.

The LENA System is intended: 1) to provide a measurement tool to help researchers gain insight into the natural language environment of children; 2) to aid professionals in the early detection of language delay; 3) to support home intervention programs directed at improving the language environment of language-delayed or disadvantaged children; and 4) to educate and provide feedback to parents regarding how much they talk to and interact with their children in order to aid them in maintaining and improving their children's language environments.

## **Abstract**

The LENA™ Digital Language Processor records real-time audio data that is transferred to a computer for processing by LENA language environment software V2.3.0. An Infoture Time Segment (ITS) file is created that summarizes information contained in the original audio file. To date, our focus has been estimating adult word counts and conversational turns between adult and key child. However, other audio factors are processed that have not yet been studied. Three versions of the LENA System exist: a consumer version; a clinician version; and a research version. The research version provides access to the ITS file. Here we describe the general content of the ITS file, detail the specific components, and discuss implications with regard to potential research goals.

## **Keywords**

Adult word count, audio-transcription, conversation, conversational turns, XML.

### 1. Introduction

The Infoture Time Segment File (ITS) is the final output from the LENA™ software analysis. One ITS file is created for every audio file that has been processed. In essence, the ITS file is an algorithmic transcription of the audio data. The ITS files are written in XML format, and are therefore versatile and user-friendly. The standard format of the XML code is easy to read and can be converted to other output formats, e.g., plain text with one segment per line. Any standard XML parser can parse the data; content in XML format can be mined intelligently to reveal new possibilities of research-oriented goals. Additionally, the XML format can be easily exported to other database files, including Excel. Data contained within the ITS file may be visualized using the LENA software V2.3.0.

### 2. Creating the ITS File

As illustrated in Figure 1, the ITS file is the final product of an iterative series of algorithmic procedures and analyses.

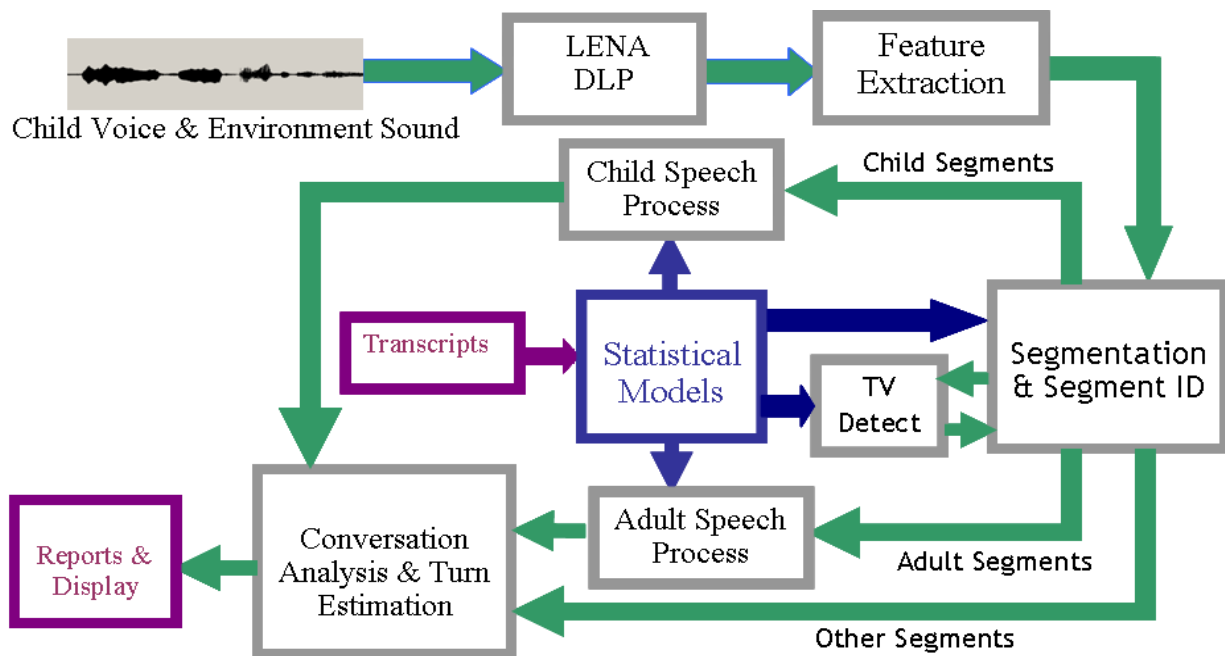


Figure 1. LENA Language Environmental Analysis Audio Processing System.

Briefly, an audio file consisting of at least 12 hours of a natural home environment audio are retrieved from the LENA digital language processor (DLP). Features are extracted from the audio data and algorithmically segmented through iterative modelling processes into first tier components consisting of key child, other child (near and distant), adult male (near and distant) and adult female (near and distant), noise, silence, overlapping sounds, and electronic (e.g. television/radio) audio categories (Table 1). Key child segments are further processed to distinguish segments containing key child speech (including words, babbles, and pre-speech communicative sounds such as squeals, growls, or raspberries) from non-speech (including cries and vegetative sounds). Adult sound segments are processed into an estimate of the number of adult-spoken words a child hears. Conversational Turns are further estimated based on all previous results. Each conversation can give such details as who is the initiator and who is the responder in a conversation. Please refer to Technical Report ITR-03-1 for further details.

Table 1: Speaker Identification Codes Assigned During Segmentation.

<b>Speaker ID</b>	<b>Description</b>
MAN/MAF	Male Adult Near/Far
FAN/FAF	Female Adult Near/Far
CHN/CHF	Key Child Near/Far
CXN/CXF	Other Child Near/Far
NON/NOF	Noise Near/Far
OLN/OLF	Overlap Near/Far
TVN/TVF	Electronic Sound Near/Far
SIL	Silence

The result of the analyses is the ITS file, a compilation of every facet of data recorded and analyzed. The ITS file can be used directly to generate reports and visual representations of the data.

### **3. General Content of the ITS File**

The DLP unit is equipped with a real-time clock (RTC) for time-stamping audio data. All dates and times are in Greenwich Mean Time (GMT) and are formatted based on ISO 8601 Date and Time Formats Conventions.<sup>1</sup> Each file contains information on the XML format (e.g. version and encoding style) and file information (e.g. name, version, and time created). An audio header section identifies recording information specific to the file. An algorithm section details the various LENA software algorithms and models used to analyze the audio data. Key child and segmentation information follow. The key child information provides key child demographics, gender, and a mean length of vocalizations. Segmentation information reveals every segment of speech spoken by adult, key child, or other child, as well as other types of non-speech sounds. Please refer to Technical Report ITR-05-1 for information on segmentation processes. Speakers were identified using codes similar to those established by professional human transcribers. Please refer to Technical Report ITR-06-1 for information on professional transcription processes. In human transcription, speakers were identified on the basis of proximity to the key child, i.e., clear or close (near-field) or distant (far-field). There are five general categories of speaker codes including human speakers (e.g. key child, adult male, adult female, or other child), unidentifiable speakers, distorted speakers, background sounds (e.g. sounds resulting from group gatherings, overlap speech segments, or silence), and other noise (e.g. television, radio, telephone, etc.). ITS File components are summarized in Table 2.

---

<sup>1</sup> Please refer to the ISO 8601 documentation for further details on time-stamp formatting (<http://www.w3.org/TR/2004/REC-xmlschema-2-20041028/#isoformats>).

Table 2: ITS File Component Description.

<b>Component</b>	<b>Description</b>
Audio Header	Recording information
Algorithm Version Section	LENA language environment analysis software models and algorithms
Key Child Information	Demographics and mean length of vocalization
Segmentation Information	Machine audio-processing result
Summary	Hourly and 5-minute summaries

Within the ITS file, every segment of a speech or a non-speech sound is classified as part of either a pause or a conversation. Pauses may contain a variety of sounds including silence, noise, overlapping speech, far-field speech, child or adult vegetative sounds, child or adult fixed signals, or electronic noise (e.g. television or radio).<sup>2</sup> The duration of a pause is by definition greater than or equal to five seconds when it occurs between two consecutive conversations. However, a pause at either the beginning or the end of a speech segment may be less than five seconds in duration.

#### **4. Research Implications**

In 2005, scientists at Infoture started a large-scale study to assess the language environment of 2-36 month-old children from families of varying socioeconomic status (SES) backgrounds. The result of this effort was a very large database of quantifiable adult-child speech phenomena in a natural home environment, known officially as the Infoture Natural Language Corpus. To date, a great deal of information has been collected

---

<sup>2</sup> Conversations may contain these same sound categories when the segment duration of the sound is less than five seconds.

into this database, specifically within the ITS files that were generated for all audio data. Thus, the potential for future data mining is considerable.

The segmentation information are conveniently summarized at the end of each ITS file. Thus, the user may easily access segmentation information at the hourly, 5-minute, or even higher resolution levels. From the segmentation summary, the user may acquire estimated values of adult word counts, conversational turns, and key child vocalizations and duration. Additional information that could be acquired includes an estimate of the number of overlapping segments, the number of interactions between key child and other child, the length of a television segment, and the amount of child vocalizations during segments containing television, conversation initiator identifications, and the number of instances of floor holding (e.g. monologue).

We have three versions of LENA, including a consumer version, a clinician version called LENA Pro, and a research version called LENA Research. The advanced LENA versions are specifically designed for pediatricians and speech language pathologists (LENA Pro) and researchers (LENA Research). The LENA Research version provides access to the ITS file under special license, and is restricted to use for qualified research purposes.

## **5. Conclusions**

The ITS file is a user-friendly, exportable XML-formatted compilation of all algorithmic analyses of the original audio data file. The file is composed of an audio header, a descriptive algorithmic region, segmentation information, and key child information. The file may be mined intelligently by the user to reveal new research possibilities, or may be used simply to report and display data that has been summarized. The content contained within the file has the potential to provide valuable information to researchers and parents alike.