Submission Cover Sheet for 1999 IEEE SoutheastCon

Submission Type (Please Check One):

☐ *Full Length Paper - (Due 11/1/98)*　　✓ *Concise Paper - (Due 12/1/98)*

Title of manuscript:
A UNIX-BASED SPEECH DATA COLLECTION PLATFORM

Authors and affiliations (List authors in the order to appear on the manuscript. List affiliations (companies, universities, etc.) and addresses exactly as they should appear. If there are more than 4 authors, please attach a second sheet.

1. Name **Richard Duncan**
Affiliation **Mississippi State University**
Address **PO Box 9571**
Mississippi State, MS 39762

3. Name
Affiliation
Address

2. Name **Dr. Joseph Picone**
Affiliation **Mississippi State University**
Address **PO Box 9571**
Mississippi State, MS 39762

4. Name
Affiliation
Address

Corresponding Author:
Name: **Richard Duncan**
Address: **PO Box 9571**
Mississippi State, MS 39762

Telephone: **601-325-8335**
FAX: **601-325-3149**
E-mail: **duncan@isip.msstate.edu**

Key Words: **speech data collection, digital telephony, application design tool**

For Technical Committee Use Only -- Do Not Write Below	
Date Received: _____	Manuscript # Assigned _____
Author Notification Date: _____	Date Camera Ready Received: _____
Decision: _____	Session: _____

Notes: _____
_____

A UNIX-BASED SPEECH DATA COLLECTION SYSTEM

Richard Duncan, Joseph Picone
Institute for Signal and Information Processing
Mississippi State University
Mississippi State, Mississippi 39762, USA
Ph (601) 325-3149 - Fax (601) 325-3149
{duncan, picone}@isip.msstate.edu

ABSTRACT

It is highly desirable to collect speech data from the telephone network via a digital interface. This avoids an additional A/D conversion normally required by analog telephone data collection hardware. A popular solution to this problem is the use of a T1 line which offers 24 digital phone lines. The leading T1 interface for Sun workstations is a system developed by Linkon Corporation. Using the Linkon framework, we have developed a fully-expandable, robust environment for platform-independent collection of telephone speech data. The object-oriented software libraries and intuitive GUI provide powerful tools with which even a novice user can efficiently prototype complex applications. The system is currently being deployed by the Linguistic Data Consortium to collect part of the next SWITCHBOARD Corpus.

A UNIX-BASED SPEECH DATA COLLECTION SYSTEM

Summary

In the last ten years a direct digital interface to the telephone network has become the standard method of speech data collection for large speech corpora involving telecommunications applications. In particular, the T1 interface is popular because it is a cost-effective way to deliver a large number of voice channels. Unfortunately, such systems, most often based on PC hardware, use closed architectures consisting of proprietary software and hardware designs. This results in a strong dependence on custom software from a single vendor. Vendors have repeatedly demonstrated an inability to deliver timely and economical solutions for speech research, resulting in a great deal of wasted time and money, with no industry-standard solution in sight. In this paper we introduce a Unix-based platform for speech data collection based on an open-architecture design and using a very inexpensive and popular workstation, a Sun Sparcstation 5, as its host.

T1 interfaces for Sun workstations are hard to find. The market leader in this niche arena is a system developed by Linkon Corporation. This system includes a T1 communications card that occupies one SBus slot and performs all data transmission functions, and a two-slot multi-DSP module that handles all call processing. The Linkon board is accompanied by some low-level, general-purpose software which serves as the foundation for building telecommunication applications. We have developed a hierarchy of software libraries extending the Linkon system to elegantly perform the interactions specific to speech data collection. We provide templates for the most common types of data collection: prompt and record modalities required for corpora such as POLYPHONE, and two-sided conference-style recording required for SWITCHBOARD-type corpora. We also provide easy to use interactive tools to build such applications from scratch.

The development effort for this data collection system involved four components: the specification of an operational T1-based system, a C++ wrapper to abstract vendor-specific hardware functions, a fully hierarchical C++ code-base to manage all aspects of a speech data collection system, and an intuitive graphical user interface (GUI) that allows rapid prototyping of data collection applications. We also conducted extensive testing on naive subjects to optimize the GUI so that minimal training is required to bring new engineers on-line with the system. The high-level data collection software (including the GUI) is hardware independent, making the future transition to other hardware systems a viable proposition.

We have developed a fully-expandable, robust system for platform-independent collection of telephone speech data. Our object-oriented software libraries and easy-to-use GUI provide a powerful tool for efficient generation of complex applications. The system is currently being deployed by the Linguistic Data Consortium to collect part of the next SWITCHBOARD Corpus. As the speech community continues to focus on SWITCHBOARD-type telephone data, this system will play a pivotal role in speech recognition research.