LVCSR-BASED LANGUAGE IDENTIFICATION

T.Schultz, I.Rogina, and A. Waibel

Interactive Systems Laboratories
University of Karlsruhe (Germany), Carnegie Mellon University (USA)
{tanja,rogina,waibel}@ira.uka.de

ABSTRACT

language identification is an important prob-
ilingual speech recognition and un-
ding a language identifica-
studied the influ-
edge sources
on

3. OVERALL SYSTEM STRUCTURE

There are several kinds of architectures for LID sys-
tems. An *integrated* architecture consists of a single
global recognition system which is language-indepen-
as described in [3]. One drawback is the increas-
ty when adding languages to be identified

tures, for each language to be
system is trained, lan-
g all systems
with

5.1. First Experiments

In earlier experiments we used German data recorded at Karlsruhe and English data recorded at CMU (to get native speakers). The CMU data are collected in office environment while the data collected at y clean. We found that testing un- tions overestimate the lan- gnificantly [9]. To sults, we

5.3. Final System

Finally we built two 4-language systems to identify
German, English, Spanish and Japanese. For these
final systems we used the new recognizer [7] which
e improved in the maintim by e.g. incorporating
the decoder and better phoneme mod-
gnizer. Therefore we called
th3LMrespec-
or-