Spoken Language Testing: Ordinate’s SET-10

TESOL 2004, Long Beach

Jared Bernstein
Ordinate Corporation
Menlo Park, California
Ordinate Technology & Services

Technology
- Automatic Scoring
- Infrastructure
- Human Scoring
- Delivery
- Data Mgt
- Telephone
- Computer
- Web

Test Service
- SET-10
- SST
- Reading
Testing Service System

Database

tests, prompts, responses

Communication Network

Delivery Interface

speech
speech
data
report

Scoring Computer
Scoring Computer
Scoring Computer

Ordinate Corporation TESOL 2004
Spoken English Tests

**SET-10**: An efficient, reliable ESL proficiency test that measures listening and speaking performance in integrated tasks.

**Junior SET (I, II)**: SET forms for early learners, with more graphic support and spoken instructions in the local language.

**SST**: Spoken Spanish Test (*others on the way)*
SET Tests

Construct: *facility in spoken English* – the ability to understand spoken English and speak appropriately in response at a native-like pace on everyday topics.

Benefits:
- Simple administration at any location
- Fair, accurate scores
- More reliable (consistent) than human judgment
- Relevant to practical language use
- Precision for formative progress tracking
**Test Format:**

Test number (PIN)

---

**Part A: reading**  8 items

**Part B: repeat Ss**  16 items

**Part C: short Qs**  24 items

**Part D: build Ss**  10 items

**Part E: open Qs**  3 items
Core Technologies

Computational Linguistics
  base materials/scoring on actual language usage

Speech Recognition
  measure what people say, and how they say it
  recognition optimized for special populations

Psychological Measurement
  careful validation assures reliability and fairness
Augmented Recognition Design

4,000,000 speech files
500,000 transcribed
250,000 human grades
> 100 native languages
Phoneme & Word Alignment

<table>
<thead>
<tr>
<th>w1</th>
<th>w2</th>
<th>w3</th>
<th>w4</th>
<th>w5</th>
<th>w6</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>p</td>
<td>ppp</td>
<td>p</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
</tr>
</tbody>
</table>

75-90 Words/Min
5.8 Phones/Sec

waveform

spectrum

words

segmentation
Machine-Human Comparison

correlation = 0.94

$N = 288$

Data from testing at US universities and companies in Europe, Asia, and Latin America.
SST ~ SPT Interview

1. SPT Interview ~ SPT ILR Estimate
2. SPT Interview ~ DLI ILR Estimate
3. SPT Interview ~ Machine score

1. Same Raters
Different Material

2. Two Rater Pairs
Different Material

3. Machine ~ Two Raters
Different Material
SST: 60-Item Sequence

17 minutes
Summary & Invitation

Automatic spoken language testing advantages: narrow construct and valid inference, cost, operational reliability, convenience.

Available now in English (for adults & children), in development for Spanish, Japanese, …

Invitation: Ordinate supports research projects in several areas – let’s test the tests.