

# SCIENTIFIC LEARNING READING ASSISTANT™: CMU SPHINX TECHNOLOGY IN A COMMERCIAL EDUCATIONAL SOFTWARE APPLICATION

Scientific  
Learning®



March 13, 2010

Valerie Beattie

Scientific Learning Corporation

# Why Speech Recognition?

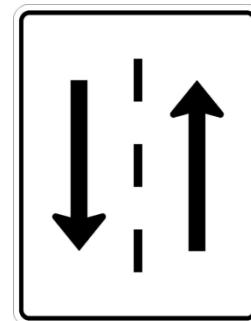
- Reading Assistant is a comprehensive software product aimed at building reading fluency, vocabulary, and comprehension
- Speech recognition technology allows the user to practice Guided Oral Reading with interactive feedback
- Guided Oral Reading practice is important for improving reading **fluency**:
  - the ability to read with sufficient ease and accuracy that the reader can focus attention on the meaning and message of text\*

*\*Jay Samuels, Reading Fluency: Its Development and Assessment, 2002 IRA What Research has to say about Reading Instruction, p.166*



# Fluency and Comprehension Closely and Causally Linked

- Research over many years has affirmed the high degree of correlation between fluency and comprehension – over 85%
- Klauza and Guthrie: “...There is evidence that fluency is both a contributor to and a product of comprehension.”



*Journal of Educational Psychology 2008, Vol. 100, No.2, 310-321*



# Why CMU Sphinx?

- Surveys of available speech recognition technologies led to the initial choice (and continued use) of Sphinx for the following reasons:
  - Source code access allowed for needed modifications, feature access, and porting
  - Sphinx-2 had important features required for incorporation into this product:
    - API
    - Real-time recognition performance
    - Acoustic model training package
  - **It's free!**



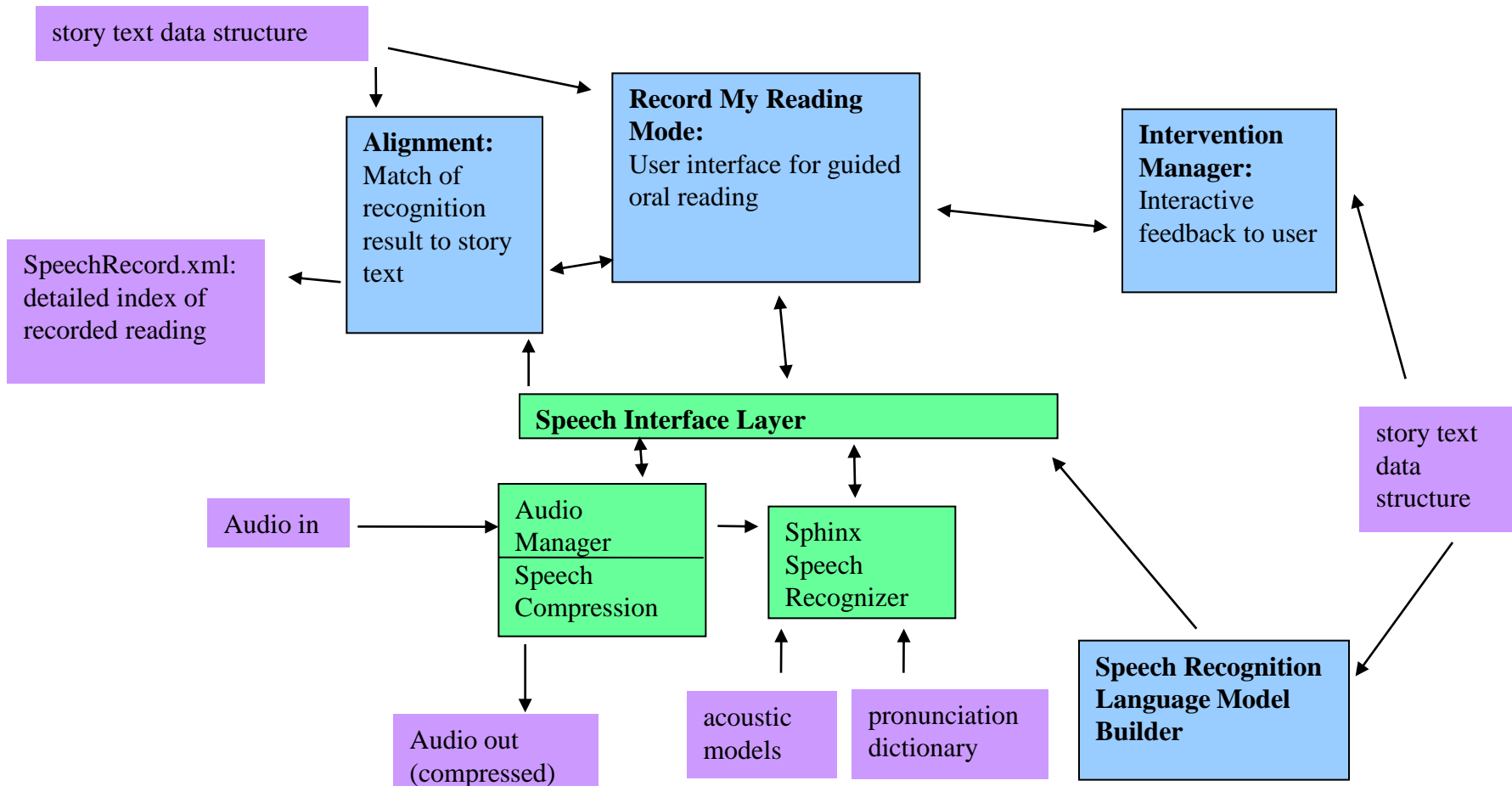
# Reading Verification

The Reading Assistant's use of speech recognition technology is different than mainstream applications of this technology:

- Typically, the goal of a speech recognition application is to determine *what* the user said.
- In the Reading Assistant application, the goal is to determine *whether* the user read the text presented, and *how well* the user read it:
  - Overall rate (Words Per Minute)
  - Particular words that the user struggled with
- This focus informs user interface design, recognizer configuration, and performance metrics



# 'Record My Reading' Architecture



# Timeline of Sphinx Development for Reading Assistant

Soliloquy Learning

Scientific Learning

	2002	2003	2004	2005	2006	2007	2008	2009
<b>Version</b>	2.0	2.1	3.0	3.2-3.3	4.0	4.1	4.1.2	Expanded Edition
<b>Recognizer</b>	Microsoft SAPI	Sphinx-2			PocketSphinx merge			
<b>Acoustic Models</b>		Children's acoustic models based on licensed corpora			Improved female voice model			Southern dialect acoustic models
<b>Recognizer Changes</b>		<ul style="list-style-type: none"> <li>Port to Mac</li> <li>Language Model 'Reading Verification' strategy developed</li> </ul>	<ul style="list-style-type: none"> <li>Word foils</li> <li>Automated testing capability</li> <li>Language Model implementation re-designed</li> </ul>		<ul style="list-style-type: none"> <li>Automatic acoustic model selection</li> <li>Vocal tract length norm.</li> </ul>		Improved automated testing	
<b>Audio Data Integrity</b>		Initial Mic Check	Continuous checking of audio input integrity					





# Beyond the Recognizer: User Interface and Interpretation

- Goal is to build fluency – UI must not interfere with that!
  - Can't be too distracting
  - Can't over-correct for minor or non-errors: 1% target for 'false negative' errors
- Alignment: Follow closely where user is in the text, even when reading or recognition has errors
  - Hierarchy of word importance
  - Partial recognition results essential – can't wait for user to pause
- Intervention: Give help in a timely, user-friendly way
  - Intervene faster, or slower, depending on reading level and user behavior
- Majority of patent applications we have submitted cover innovations in this area





# DEMO

