

The use of NLP technologies to engineer oral proficiency test items

Ross Hendrickson, Deryle Lonsdale
ross.hendrickson@gmail.com, lonz@byu.edu
Brigham Young University



Problem : Testing EI

- Elicited Imitation
- Hear and repeat
- Requires in-depth item creation
 - The husband has become a father.
 - Big ships will always make noise.
 - He had a voracious appetite the height of a towering weed and the strength of a malicious rancorous hippopotamus.

Features: What we are looking for and why

- Previous work analyzing influence of sentence-level features on an IRT based difficulty score using Akaike's Information Criterion and stepwise linear regression.
- Features were chosen that were found to be significant for different syllable counts.
- Examples of some of the 23 features recognized by the tool
 - Past tense
 - Lexical Frequency
 - Syllable Count
 - Simple or Complex Sentence

What the tool can do

- Sentence level feature annotation
- Parse documents and annotate each sentence
- Interface with extant Penn Treebank databases and annotate them
- Retrieve sentences with desired features from database

Incorporated NLP tools

- Wordnet 3.0
- Stanford Parser
- Stanford Tregex
- CELEX v2.0
- British National Corpus

Code Snippets

- Sample Tree - Regular expressions
 - @DT
 - SBAR <, IN|WHADVP|WHNP|WHVP
- Wordnet – Semantic Calculation
 - In 7 another 1 moment 6 down 26 went 30 Alice after 4 it 1 never 2 once 3 considering 9 howin the world 9 she was 14 to get 37 out 17 again 1 .
 - Avg Senses: 11.133

Sentence Example 1

- Simple Sentence
- Low Syllable Count
- Past Tense
- Determiners
- Declarative Sentence

A player threw
the ball to me.

The screenshot shows a software interface with a text input field containing the sentence "A player threw the ball to me." Below the input field are a "send to DB" checkbox and a "Parse" button. The "Results:" section displays several statistics: Syllable Count (8.0), Morphological Calc (1.2), Semantic Calc (10.0), Syntactic Calc (19.0), Average raw lexical frequency (1550266.0), and detected categories (Articles, Past Tense). The parse tree below shows the hierarchical structure of the sentence, starting from a ROOT node which branches into S. S branches into NP and VP. The first NP branches into DT ("a") and NN ("player"). VP branches into VBD ("threw"), NP, and PP. The second NP branches into DT ("the") and NN ("ball"). PP branches into TO ("to") and NP, which further branches into PRP ("me").

```
graph TD
    ROOT --> S
    S --> NP1[NP]
    S --> VP
    NP1 --> DT1[DT]
    NP1 --> NN1[NN]
    DT1 --> a[a]
    NN1 --> player[player]
    VP --> VBD[VBD]
    VP --> NP2[NP]
    VP --> PP[PP]
    VBD --> threw[threw]
    NP2 --> DT2[DT]
    NP2 --> NN2[NN]
    DT2 --> the[the]
    NN2 --> ball[ball]
    PP --> TO[TO]
    PP --> NP3[NP]
    TO --> to[to]
    NP3 --> PRP[PRP]
    PRP --> me[me]
```

Sentence Example 2

- Complex Sentence
- High Syllable Count
- Past Tense
- Possesive
- Morphologically simple
- Semantically Complex

Hesitating before she spoke her next line the actress reached the pinnacle of her nervousness.

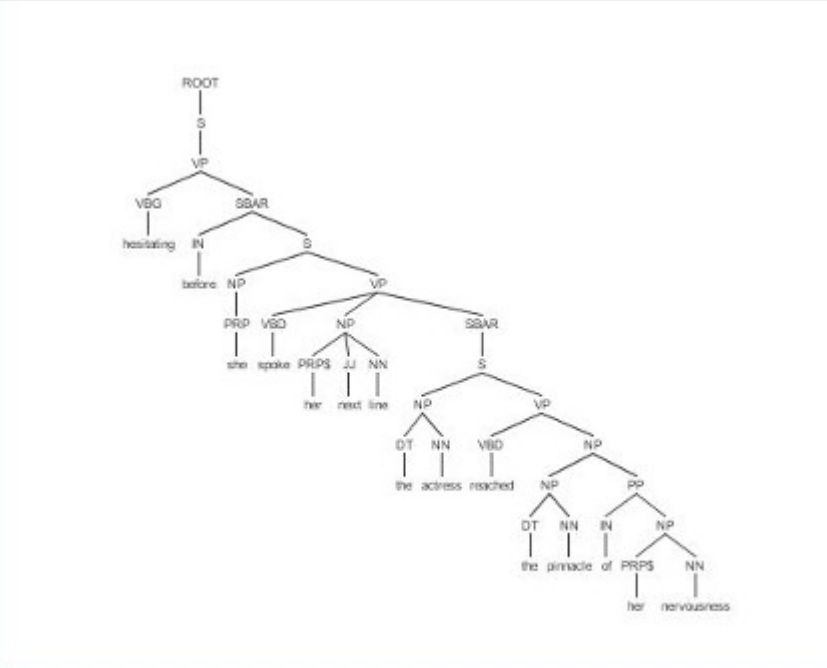
Sentence Document Corpus Generation

Hesitating before she spoke her next line the actress reached the pinnacle of her r

send to DB

Results:

Found: Syllable Count : 24.0
Found: Morphological Calc : 1.1538461538461537
Found: Semantic Calc : 8.0
Found: Syntactic Calc : 30.0
Found: Average raw lexical frequency : 1084483.0
Found: Articles
Found: Past Tense



Documents

- Typical News article (approx 40 sentences)

After getting word that the long rumored *The Beatles: Rock Band* game would be coming later this year we turned our attention to the existing *Rock Band* music. There have been two versions of the game released in *Rock Band* and *Rock Band 2* as well as a ton of full of music and hundreds of songs via DLC.

As of our research, the official tally of playable music in the *Rock Band* catalog is 17 genres. We wanted to take a look at this list and see what we can tell about the songs that have been included so far. We took time to make graphs to show some of the data compared by decade, source, and more. Some of what we found is a little surprising.

WHERE IS THE MUSIC FROM?

One of the best features of *Rock Band* has been the compatibility of the purchases from standalone games came packed including one devoted entirely to downloadable content available via download as either

Both Xbox Live and Playstation Network have seen the music get added within days of each other. The downloadable content. The downloadable weekly DLC updates of back music getting.

Right now, the Wii has 27% of the music have been made available so far

```
input - Notepad
File Edit Format View Help

What kind of Music is in this Catalog?
There are 17 genres represented in the catalog so far. It is worth pointing out that
Most of the music in the catalog are the master recordings from the original
The graph shows that the music is heavily dominated by rock, not surprising
Rock Band Music Compared by Genre

What Decades Feature the Most Music?
RB has music dating back to the 1960's and from each decade since. There is
Almost 40% of the music that is playable is from the last 9 years, which is a
```

Sentence Document **Corpus** Generation

File Path

Statistics:

```
-----
Sentences: 43
Words: 892
-----
```

: Celebrity women that Play V
an or Megan Fox from the list,

the "hard" songs. we took the
the Rock Band competition being

Extant Corpus Integration

- Input follows Penn Tree Bank WSJ format

The screenshot shows a Windows file explorer window at the top, displaying the directory E:\WSJ\00. Below it is a software interface with a 'Path' field containing 'E:\WSJ\' and a 'File Type' dropdown set to 'MRG'. A 'Parse' button is visible. The 'Statistics' section shows the following data:

Statistic	Value
Files	499
Sentences	49,208
Words	1,006,459

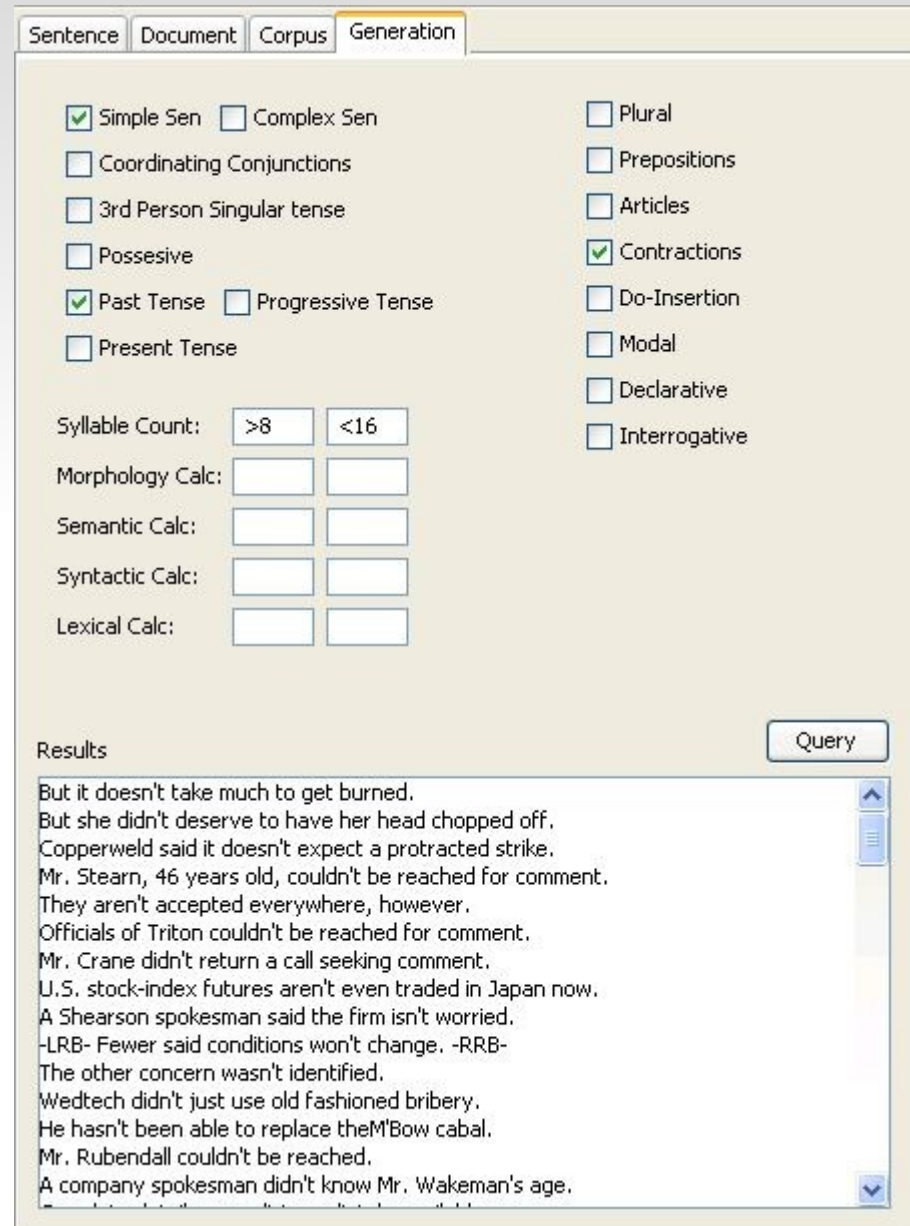
The screenshot shows a text editor window titled 'WSJ_0001.MRG' containing a parse tree for the sentence 'Pierre and Vincken are 61 years old and will join the board as a nonexecutive director of Nov. 29'. The tree structure is as follows:

```
( (S
  (NP-SBJ
    (NP (NNP Pierre) (NNP Vincken) )
    ( , , )
    (ADJP
      (NP (CD 61) (NNS years) )
      (JJ old) )
      ( , , ) )
    (VP (MD will)
      (VP (VB join)
        (NP (DT the) (NN board) )
        (PP-CLR (IN as)
          (NP (DT a) (JJ nonexecutive) (NN director) ))
          (NP-TMP (NNP Nov.) (CD 29) )))
      ( . . ) ) ) )
```

Simple Sentence Retrieval

- Simple Sentence
- Past Tense
- Contractions
- Syllable Count between 8 and 16

But it doesn't take
much to get
burned.



Sentence Document Corpus **Generation**

Simple Sen Complex Sen Plural
 Coordinating Conjunctions Prepositions
 3rd Person Singular tense Articles
 Possessive Contractions
 Past Tense Progressive Tense Do-Insertion
 Present Tense Modal
 Declarative
 Interrogative

Syllable Count:

Morphology Calc:

Semantic Calc:

Syntactic Calc:

Lexical Calc:

Query

Results

But it doesn't take much to get burned.
But she didn't deserve to have her head chopped off.
Copperweld said it doesn't expect a protracted strike.
Mr. Stearn, 46 years old, couldn't be reached for comment.
They aren't accepted everywhere, however.
Officials of Triton couldn't be reached for comment.
Mr. Crane didn't return a call seeking comment.
U.S. stock-index futures aren't even traded in Japan now.
A Shearson spokesman said the firm isn't worried.
-LRB- Fewer said conditions won't change. -RRB-
The other concern wasn't identified.
Wedtech didn't just use old fashioned bribery.
He hasn't been able to replace theM'Bow cabal.
Mr. Rubendall couldn't be reached.
A company spokesman didn't know Mr. Wakeman's age.

Complex Sentence Retrieval

- Complex Sentence
- Progressive Tense
- Plural
- Modal
- Interrogative
- Semantically complex
- Syntactically complex

Who knows what will happen down the road, in three to six months, if foreign investment starts to erode?

Sentence Document Corpus **Generation**

Simple Sen Complex Sen Plural

Coordinating Conjunctions Prepositions

3rd Person Singular tense Articles

Possessive Contractions

Past Tense Progressive Tense Do-Insertion

Present Tense Modal Declarative

Interrogative

Syllable Count: >10

Morphology Calc:

Semantic Calc: >6

Syntactic Calc: >20

Lexical Calc:

Results

Who knows what will happen down the road, in three to six months, if foreign investment starts to erode?
You have... raised important questions which ought to be answered ; What does I
Managers and presenters insist that chamber music concerts are a hard sell, but can
The precedent having been set, who can complain if future generations called upon t
I know when dividends are due ; When bonds should be retired ; But what gets by m
I expect to see some selling, but will they kill this one as they have every rally in tl
Who can make the better decision, the guy who has 10 seconds to decide what tc
The pertinent question, coming at the end of a lengthy, confidential questionnaire, w
But wouldn't a president who acted despite Senate objections be taking grave politic
As members of the Black Caucus in Congress asked during the debate on the legislati
Do the Japanese feel so superior that they can not find capable American managers?
Why should you beat your brains out fighting the environmentalists, the neighborl
If I were to ask a sample of American parents, `` Do you wish the elementary school
What kind of initiative should anyone expect from people out on the line who've read

Potential Applications

- EI item development
- Develop items for more traditional tests
- Reading comprehension tests
- ESL teachers could use the tool to develop sentences that illustrate usage of particular lexical or grammatical features being taught
- Annotate learner written language for specific sentence-level features.
- Annotate content for students to have reference guide of features / ready explanation of concepts illustrated.

Tool Impact

- Time per item generated
 - Test form E – 6 weeks (2 graduate students)
 - Test form F – 3 weeks (3 undergraduate students)
- Will significantly speed up analysis of future tests
- Provide consistent annotation

Future work

- More theory-centric calculations
 - Semantic
 - Syntactic
 - Morphological
 - Lexical Frequency
- More stringently evaluate algorithm performance
- Add additional sentence-level features of interest
- Develop word- and syllable-level features
- Possible use of database with adaptive language assessment
- Expand database selections
- Improve GUI ease of use and information display capability