

**Towards a Model for Understanding the
Connotative Meaning of Text:
A Blue Sky Project**

Liz Liddy, Mike D'Eredita, Ozgur Yilmazel

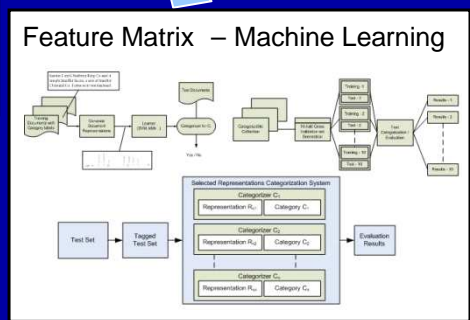
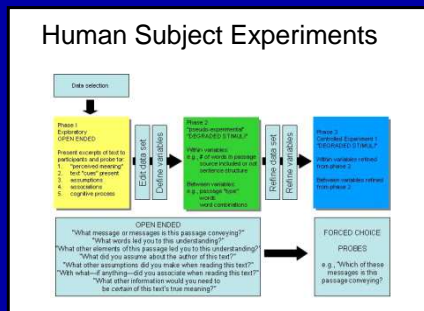
**Center for Natural Language Processing
School of Information Studies
Syracuse University**



Understanding the Connotative Meaning of Text

Syracuse University

Liz Liddy, Mike D'Eredita, Ozgur Yilmazel



GOALS

- Understand how humans extract connotative meaning from text.
 - A sincere apology*
 - An urgent request for help*
- Evaluate feasibility of a system understanding what a text suggests beyond what is explicitly stated.
- Crucial for QA systems to utilize non-specifiable content as a key dimension in interpreting how a question or answer is to be understood

PLANS

- Conduct human subject experiments to understand how connotation is conveyed & recognized.
- Integrate human subject experimental results with our NLP-based feature identification ability.
- Determine combination of features for Machine Learning that are most accurate in recognizing connotative understanding of text.
- Evaluate capability on new collection of texts.

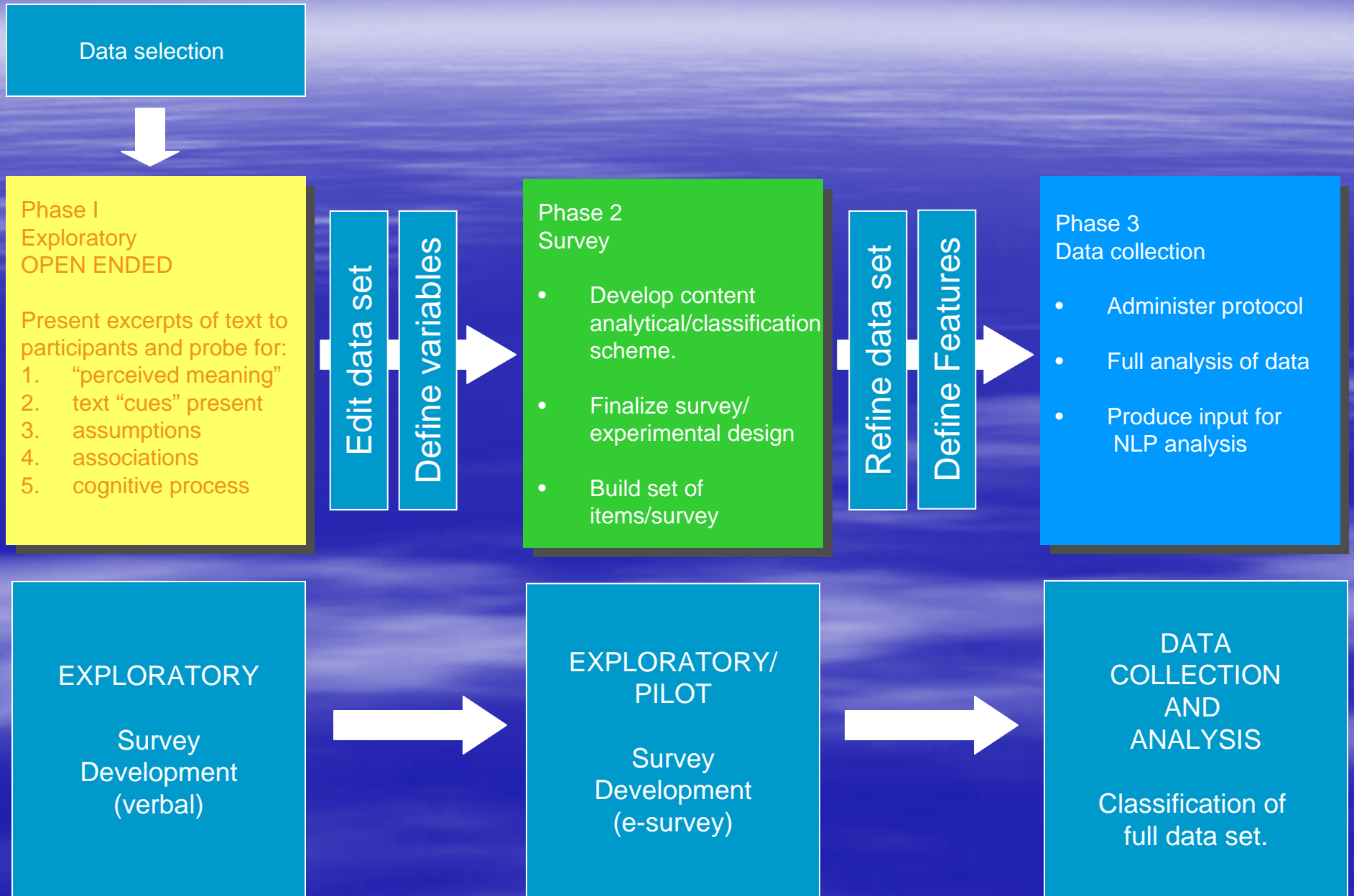
MILESTONES

- Phase I:**
 - Test case development
 - Human Subject Experiments
 - Milestone: Model Development (June 2007)**
- Phase II:**
 - Feature Matrix Development
 - Machine Learning Experimentation
 - ESM testing
 - Milestone: Connotation Detector (August 2007)**
- Phase III:**
 - Internal Evaluation
 - Milestone: Evaluation Results (September 2007)**

Theoretical Background – Linguistics

- Connotation – additional semantic interpretation of a linguistic expression beyond its literal meaning
- Integrates understanding from multiple levels of language
 - Working on specifying how we get this additional meaning
 - What features might a system utilize?
- Seeking author-independent features for distinguishing implicit connotation across sets of texts
 - First genre - blogs
 - Acquired Live Journal blog data from Mark Goldberg at RPI

Human Subject Experimental Design

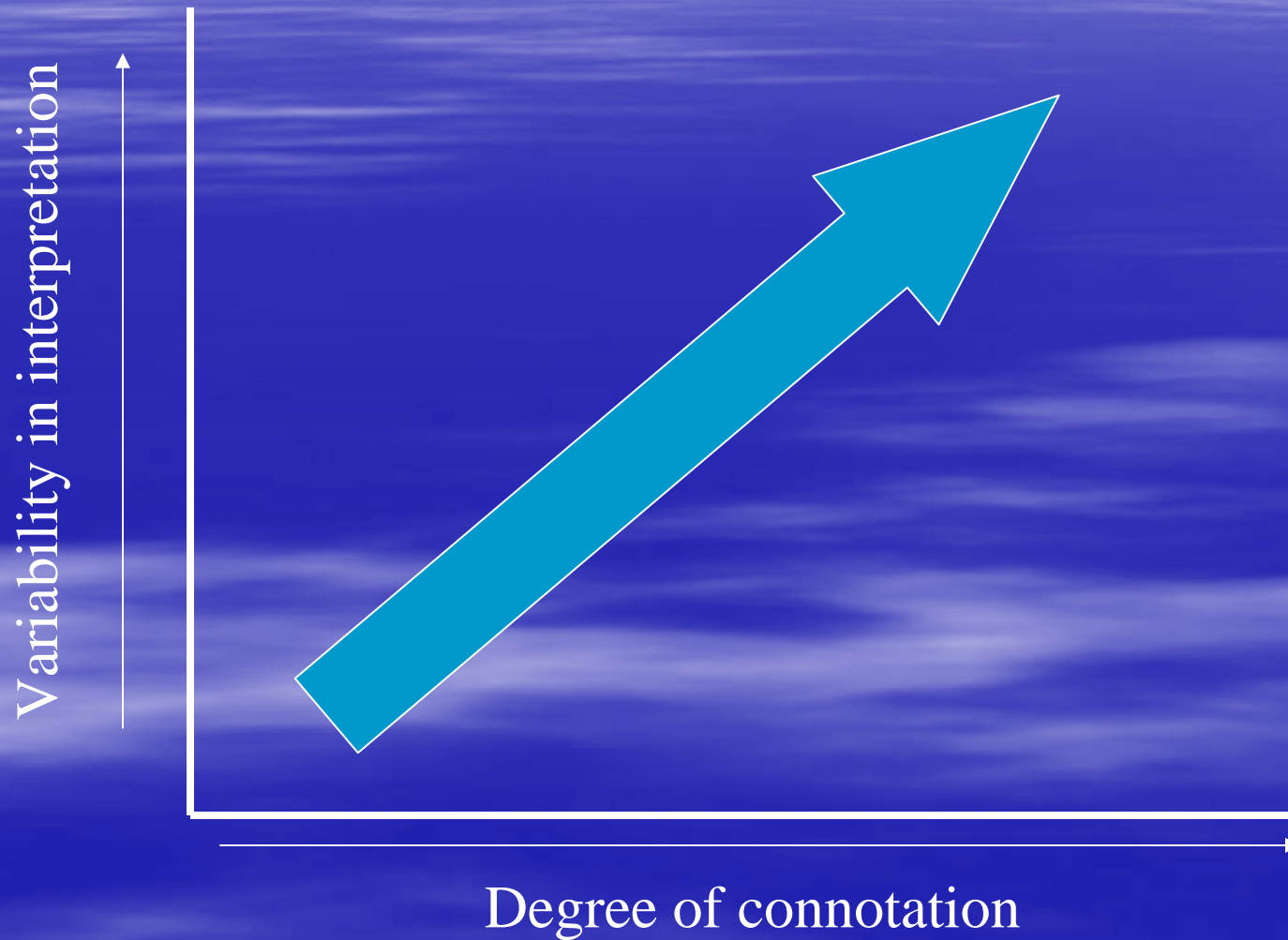


Exploratory Path To Date

- Level of analysis: word, sentence, paragraph?
- Find some data!
- Qualitatively distinguish between denotative and connotative examples
- Structured questioning with participants
 - Is this reliable?
 - 7/10 100% agreement
 - 3/10 75% agreement
 - Is this valid?
 - Degrees of connotation
- Open ended survey and content analysis

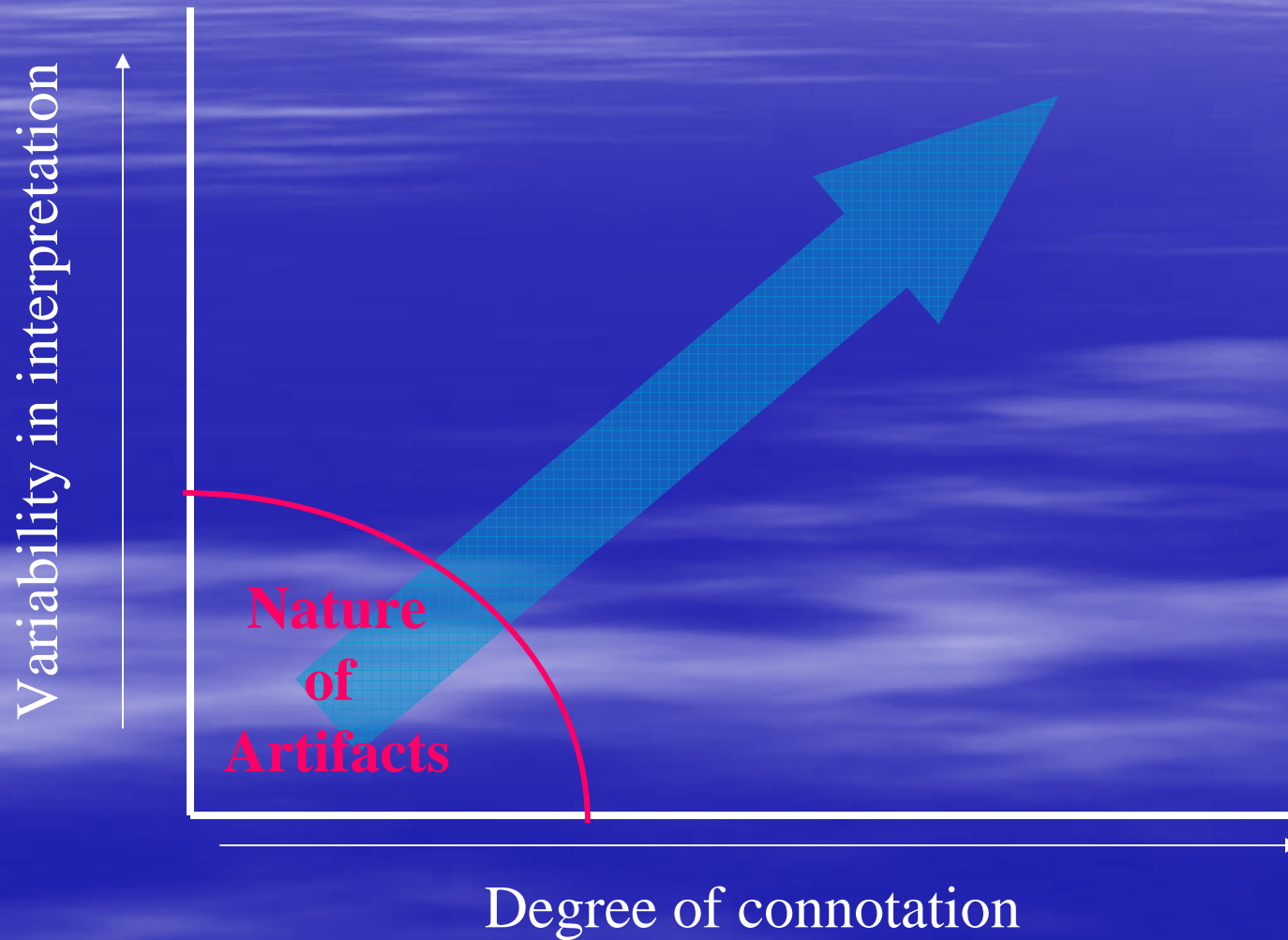
Degree of Connotation

A notion derived from data



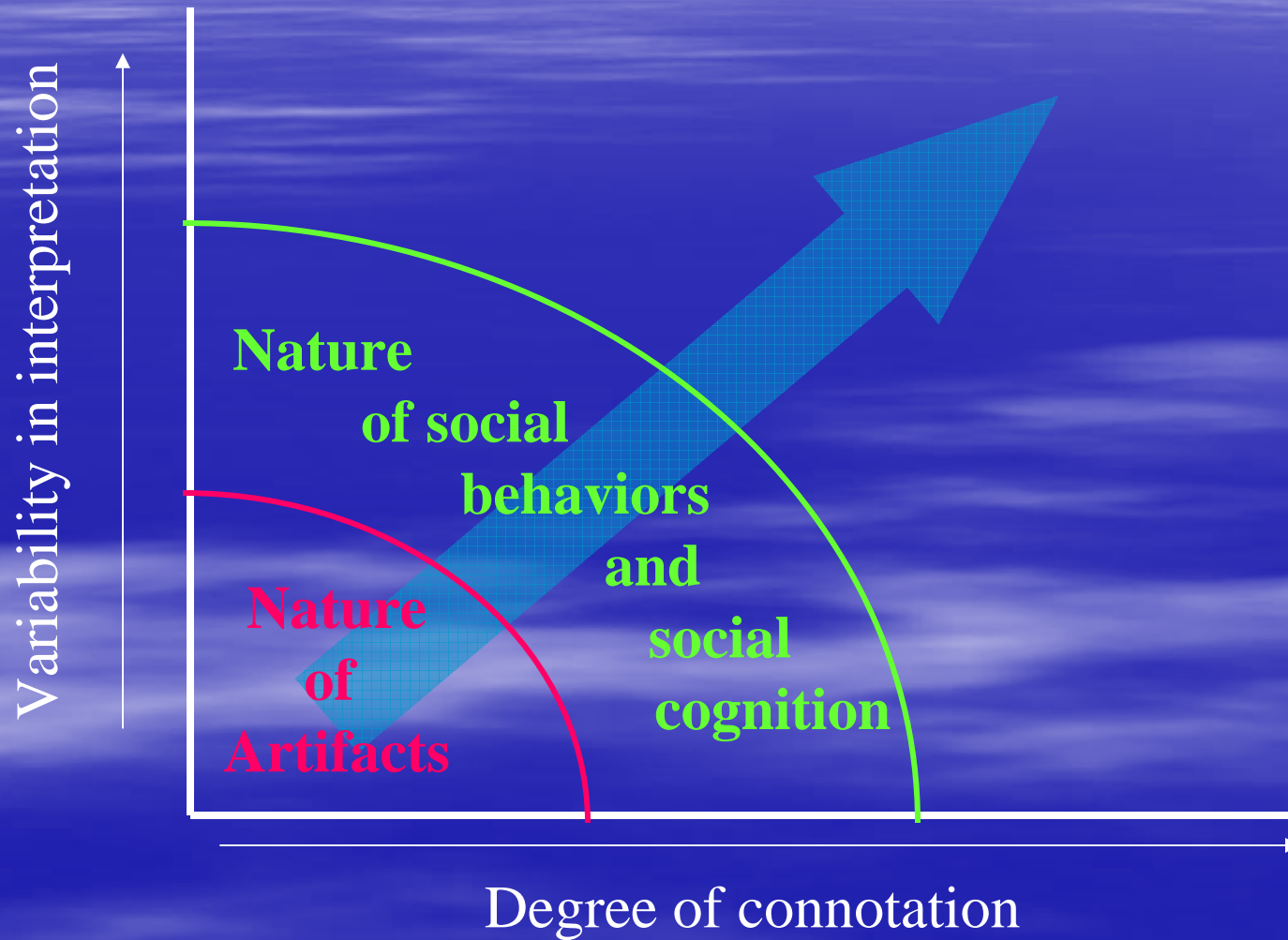
Degree of Connotation

A notion derived from data



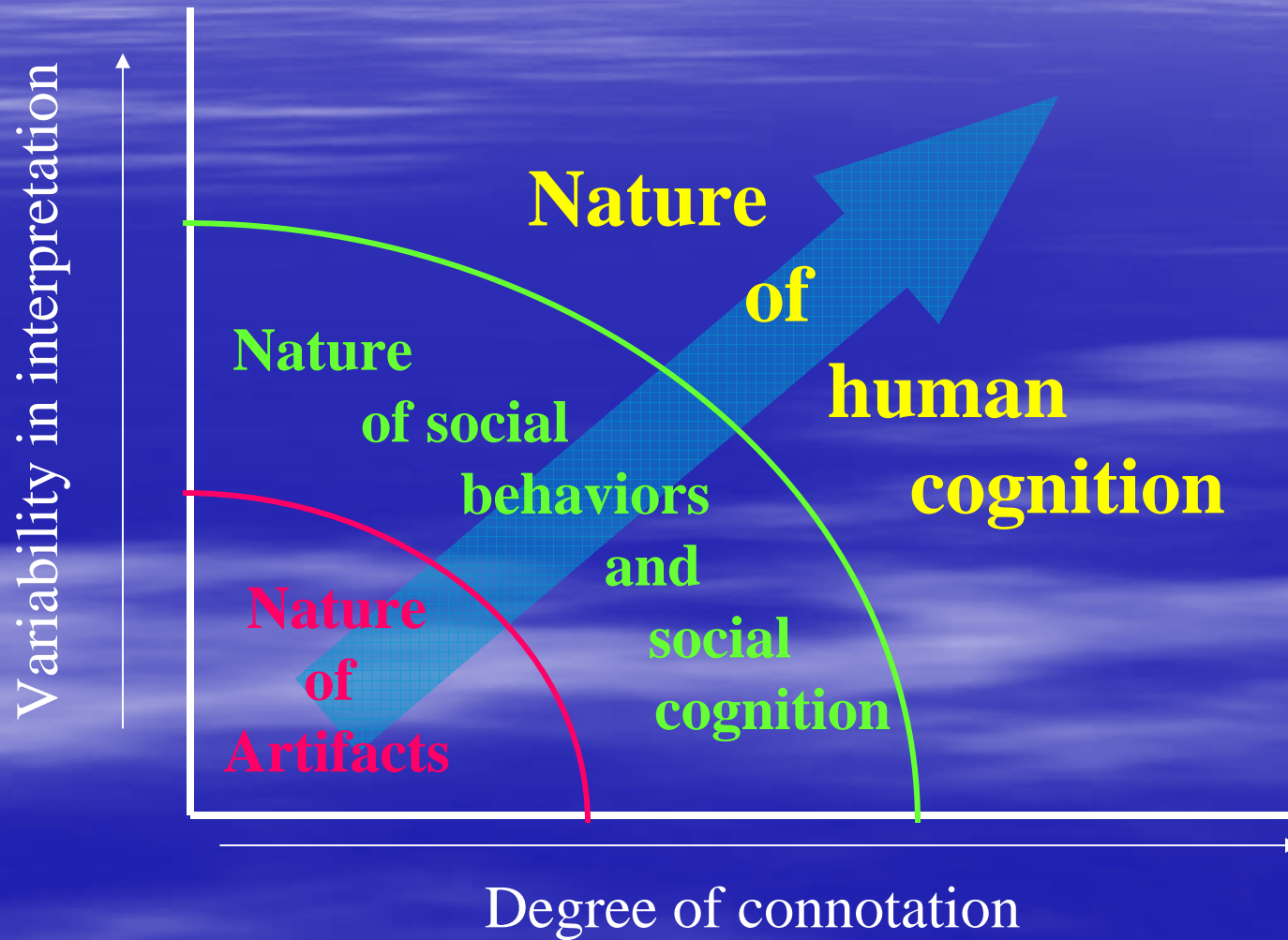
Degree of Connotation

A notion derived from data



Degree of Connotation

A notion derived from data

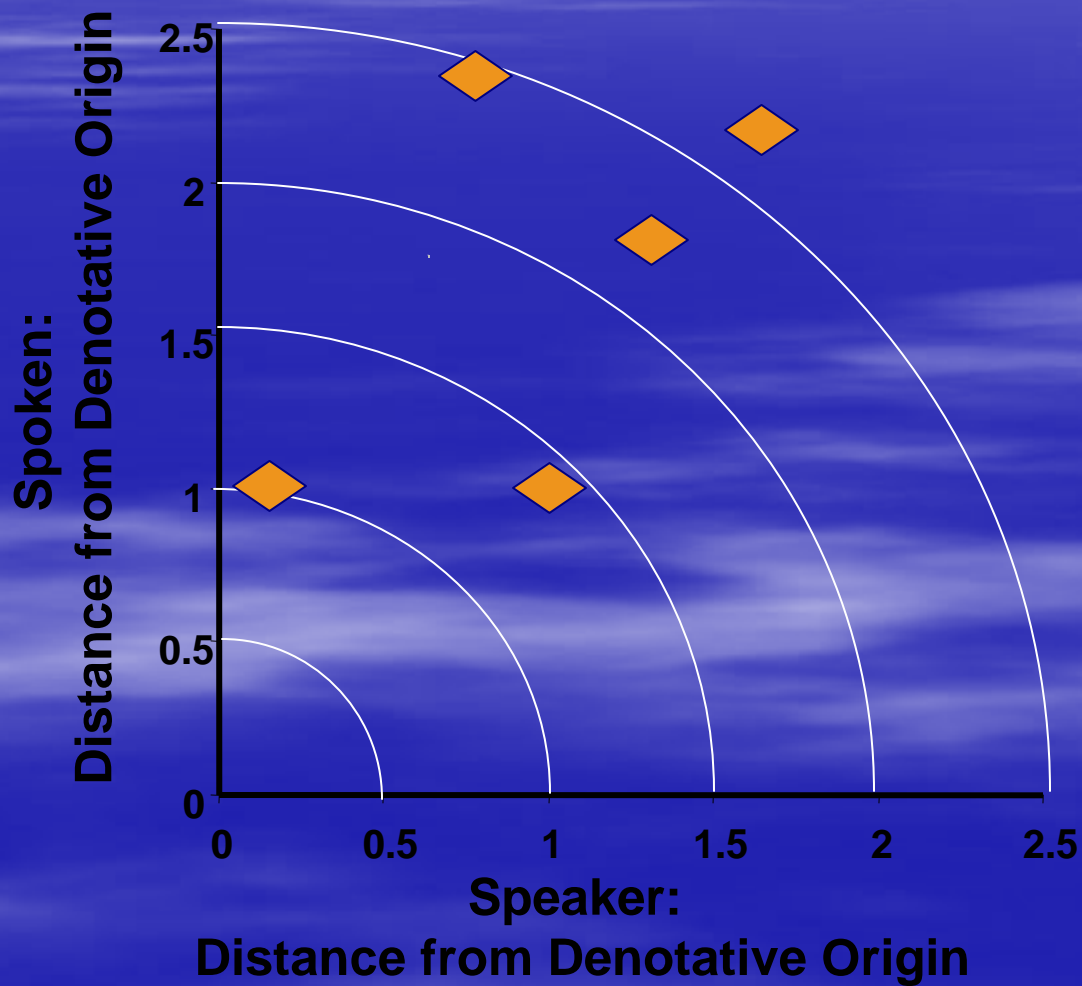


Most Current Method

- Survey 20 sentences
 - “What does this sentence suggest?” & “Why?”
 - Demographics
- Data revealed at least two distinctive types of interpretations
 - Speaker (the author)
 - Spoken
 - Message
 - Artifact
- Rough content analysis of data
 - Rate degree of connotation denoted from each response
 - 0 – 4 scale (denotative to connotative)
 - 0 denotative
 - 1 more text-based
 - 2 – 3 genre, practice-based, etc.
 - 3 – 4 individual/personal experience

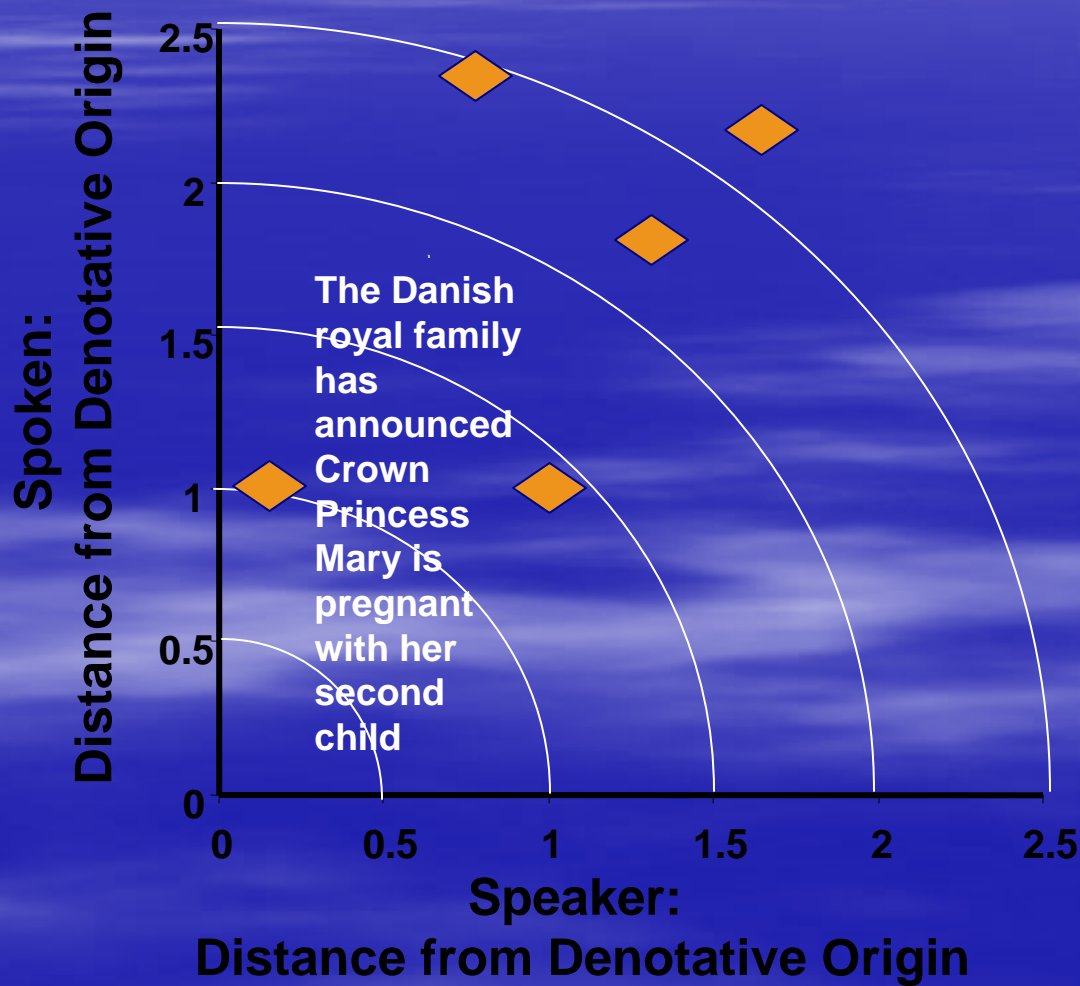
Degree of Connotation

Distance from the Denotative Origin



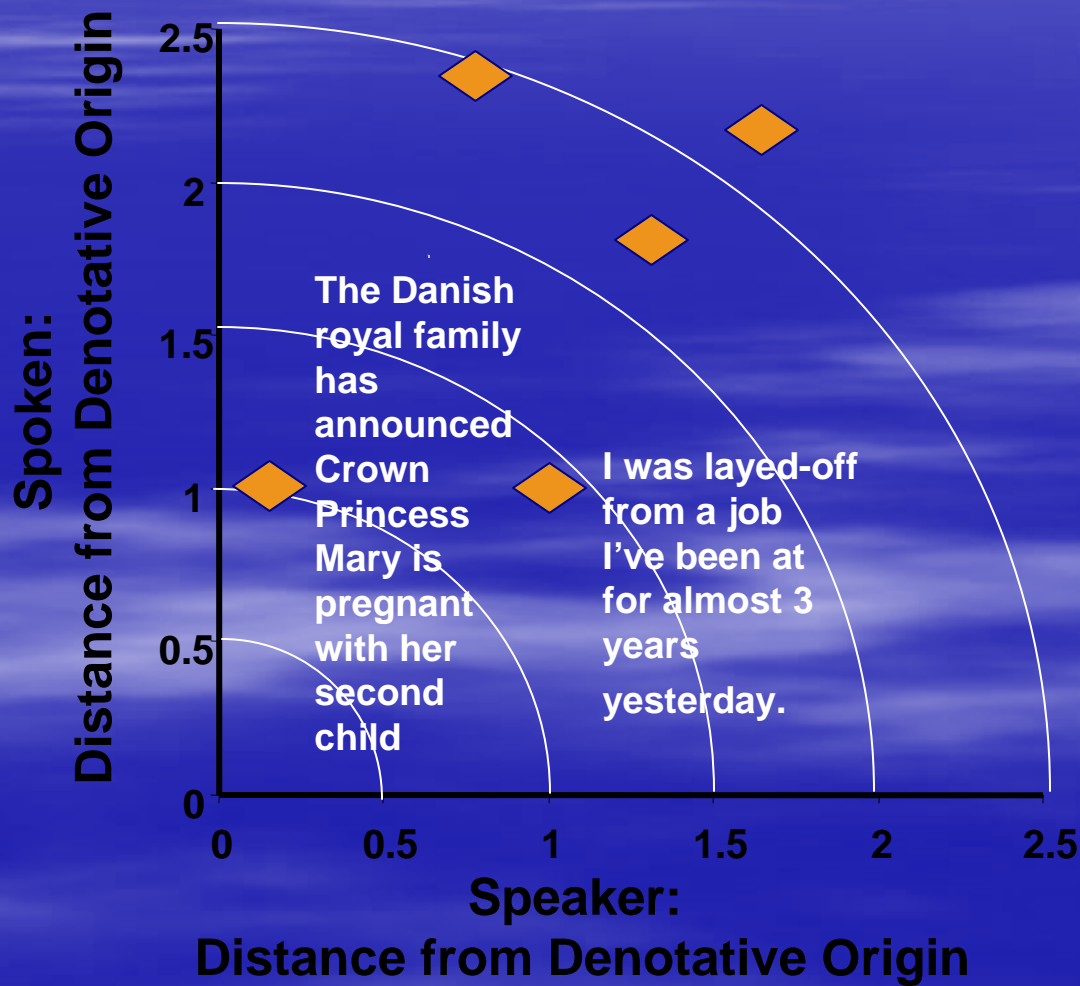
Degree of Connotation

Distance from the Denotative Origin



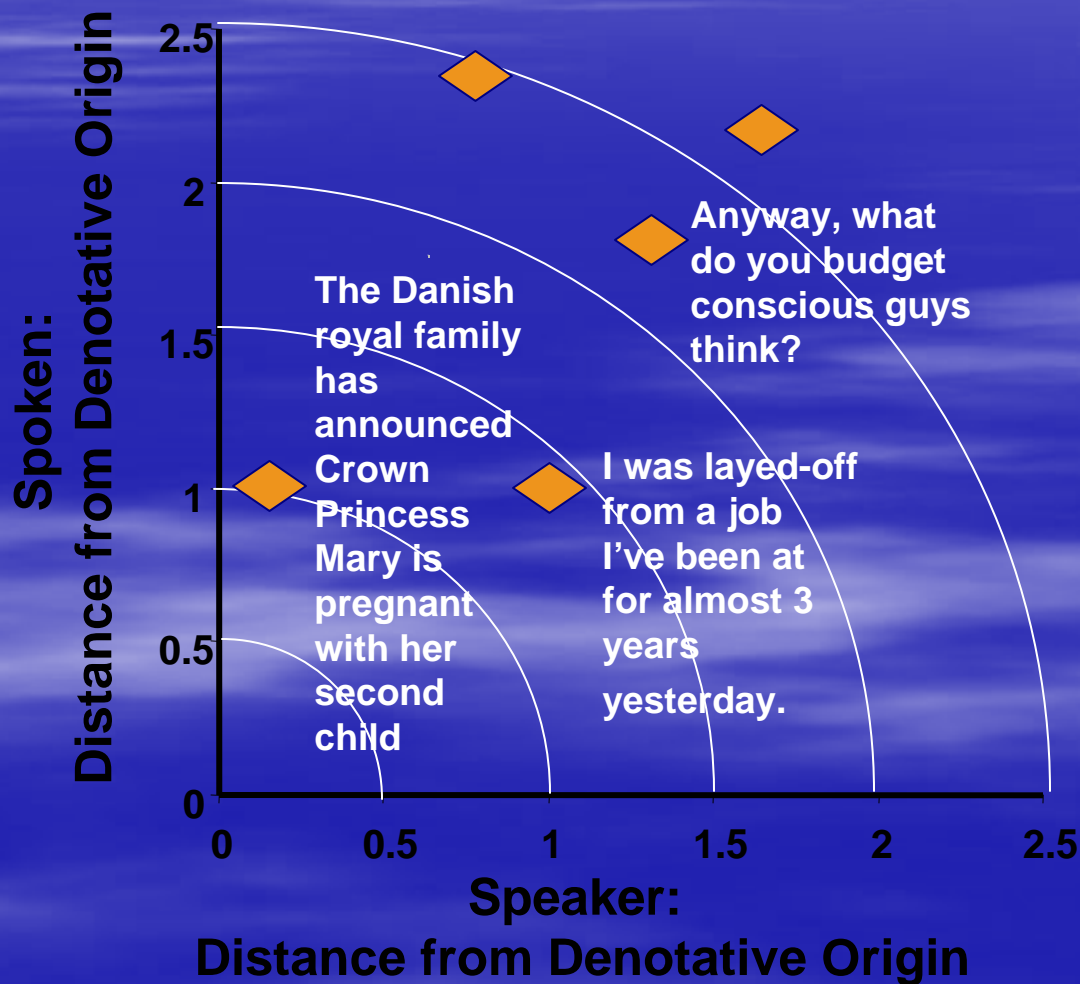
Degree of Connotation

Distance from the Denotative Origin



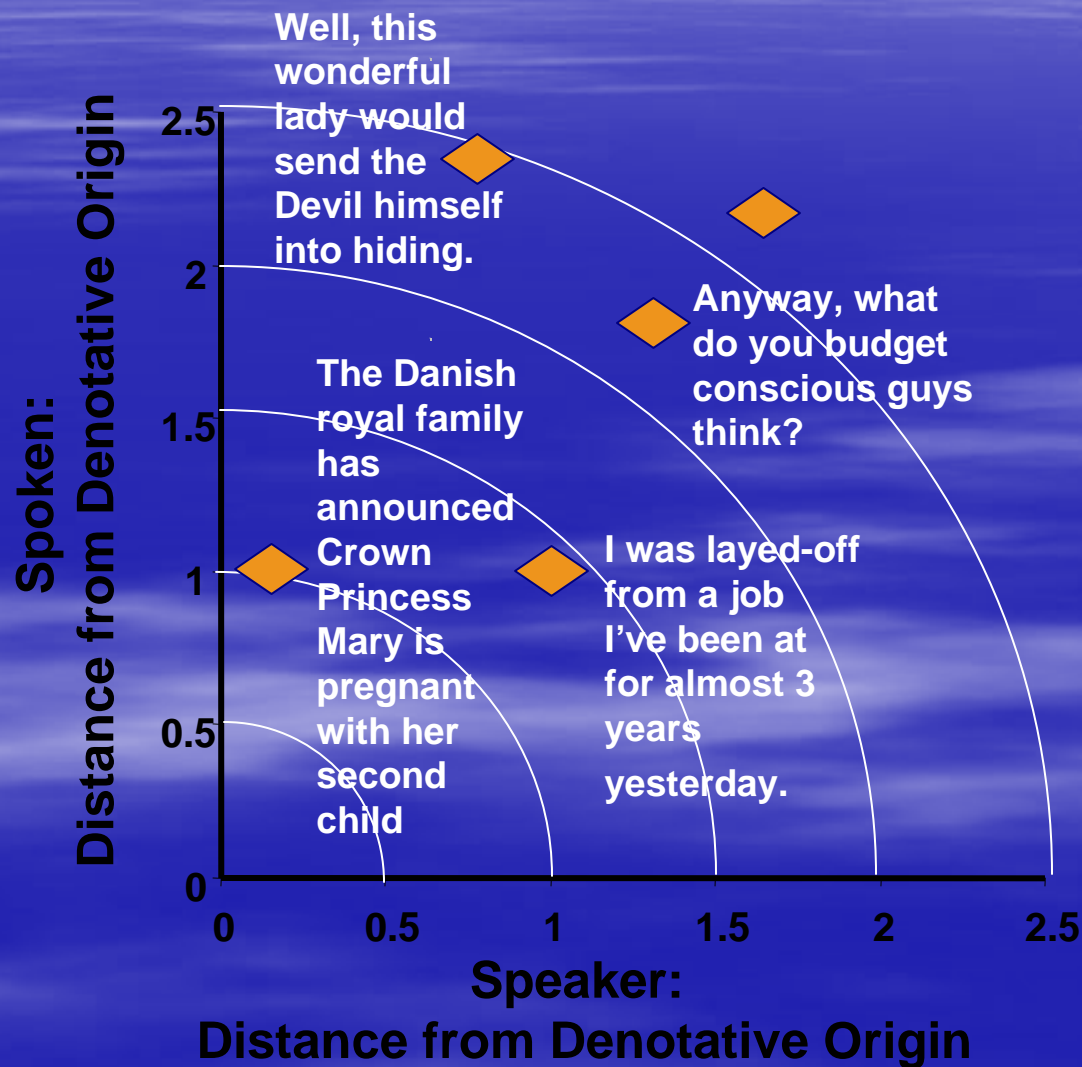
Degree of Connotation

Distance from the Denotative Origin



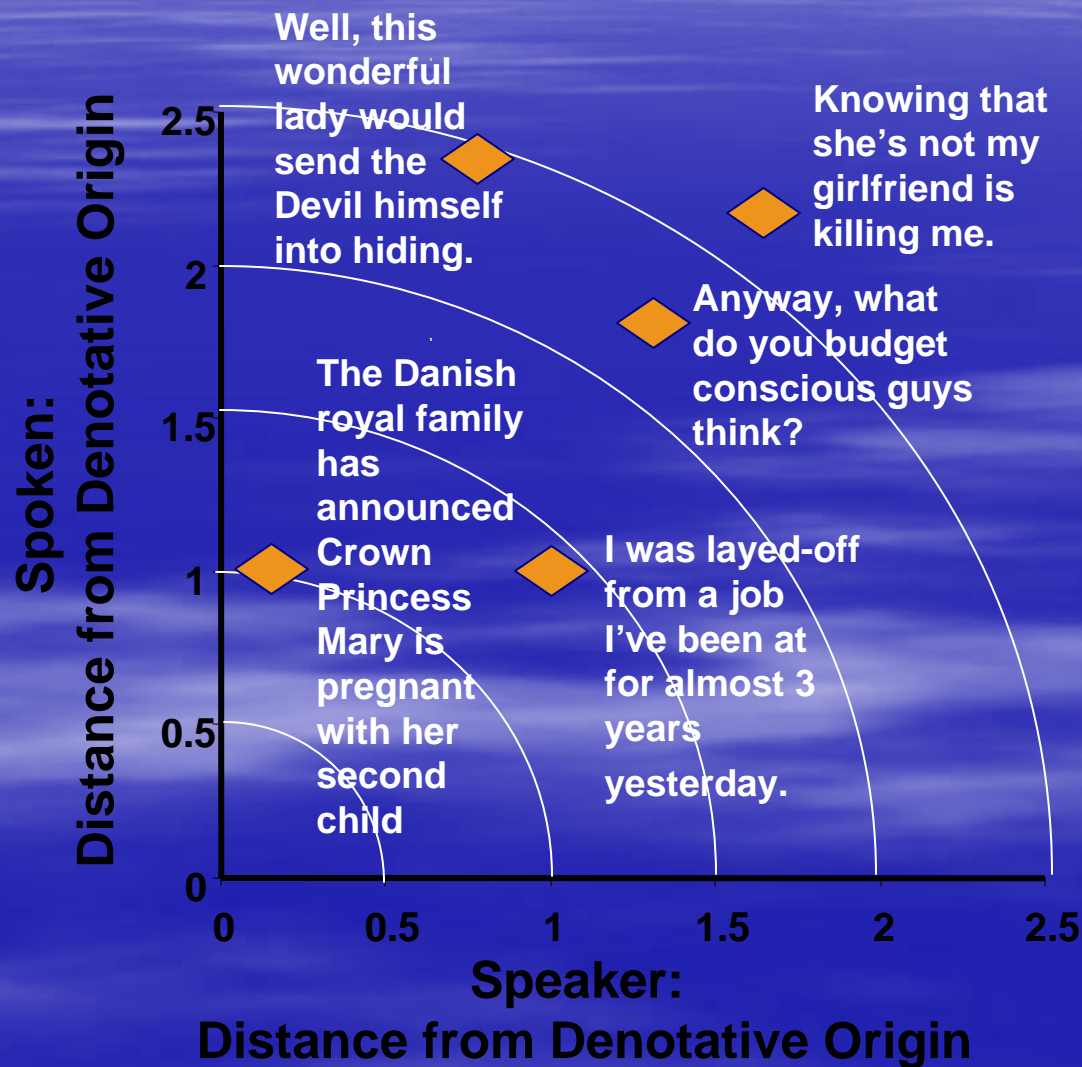
Degree of Connotation

Distance from the Denotative Origin



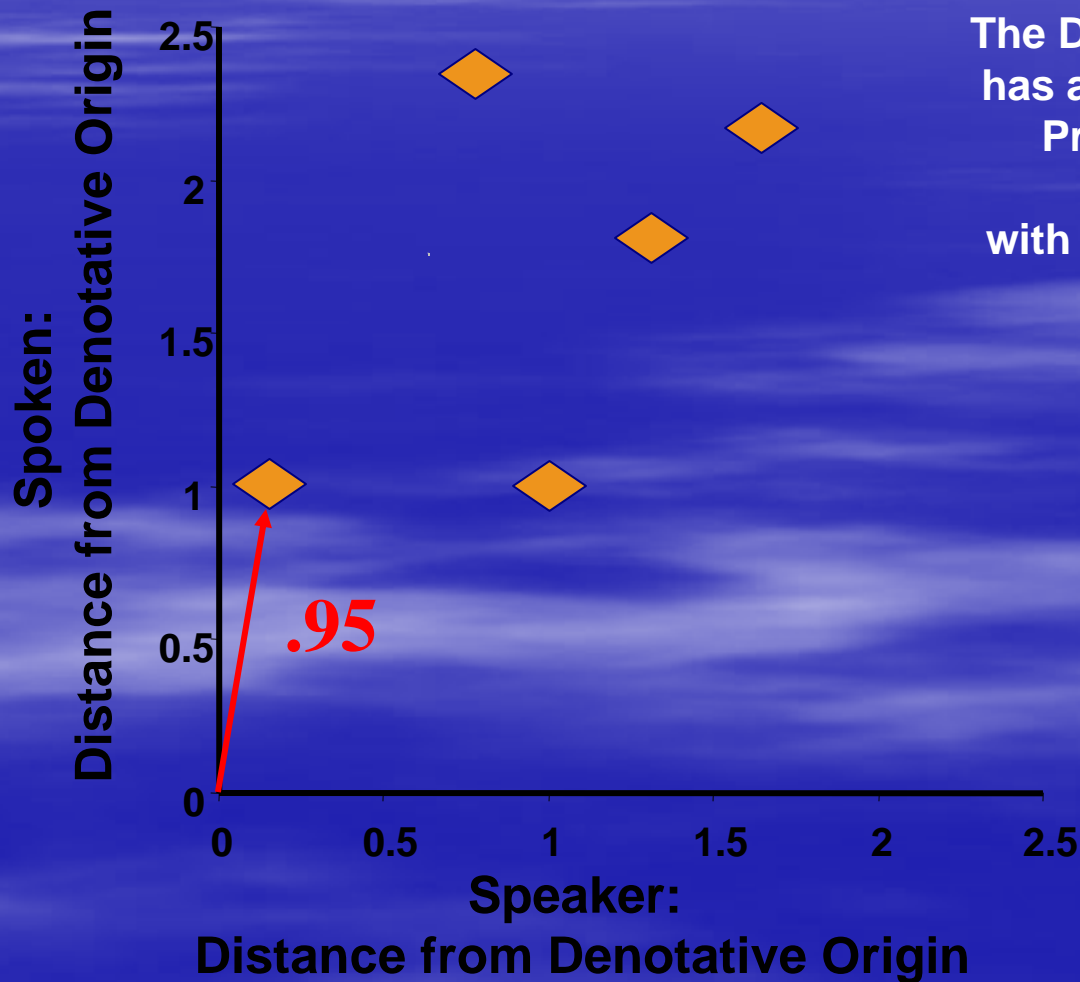
Degree of Connotation

Distance from the Denotative Origin



Degree of Connotation

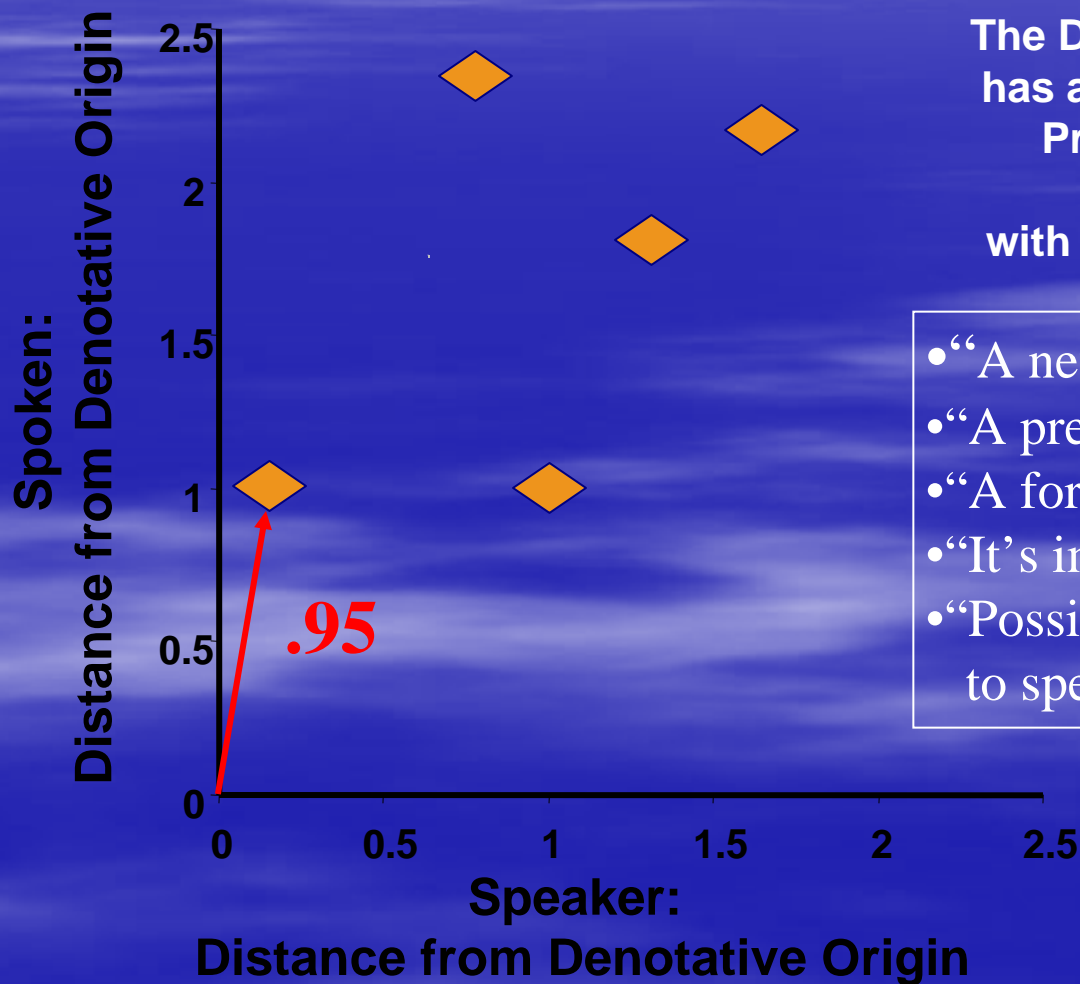
Distance from the Denotative Origin



The Danish royal family has announced Crown Princess Mary is pregnant with her second child.

Degree of Connotation

Distance from the Denotative Origin

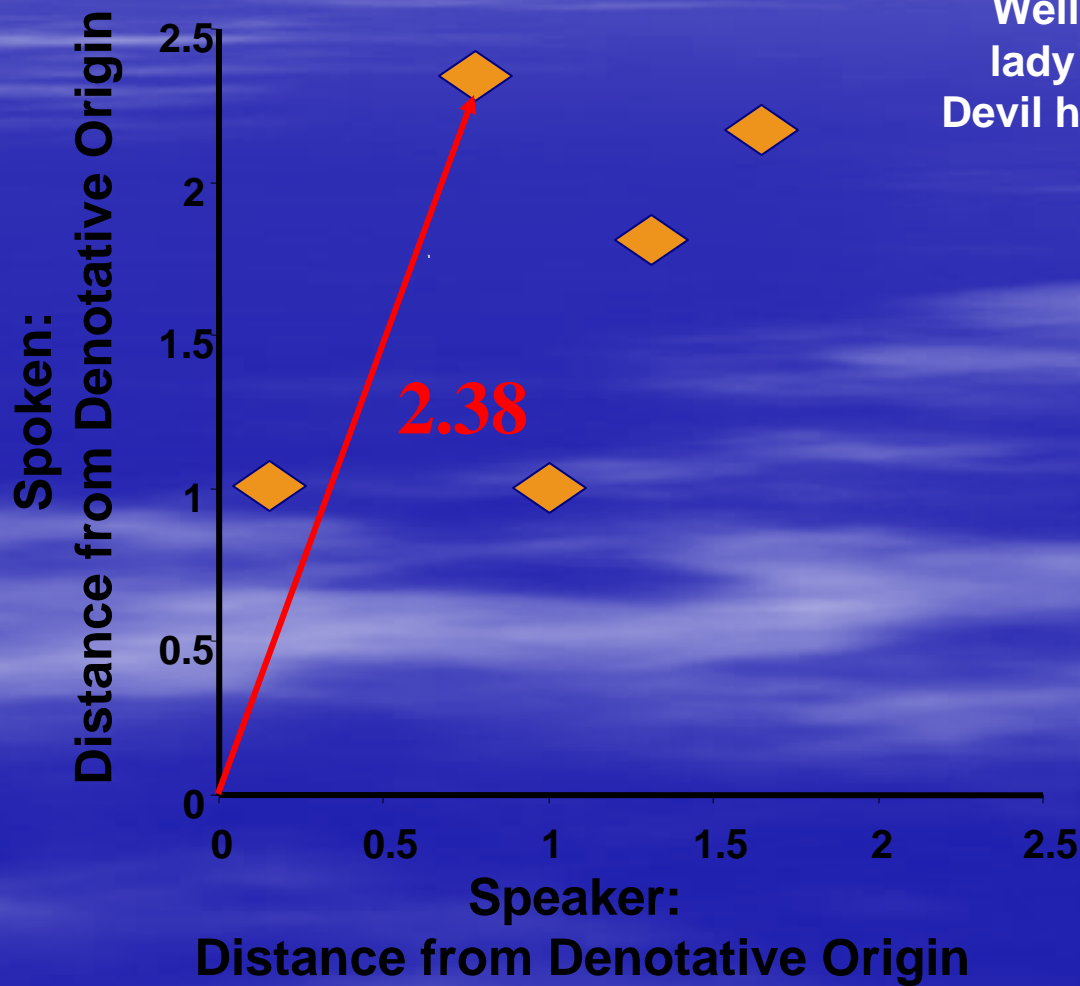


The Danish royal family has announced Crown Princess Mary is pregnant with her second child.

- “A newspaper story”
- “A press release”
- “A formal announcement”
- “It’s important news”
- “Possibly in response to speculation”

Degree of Connotation

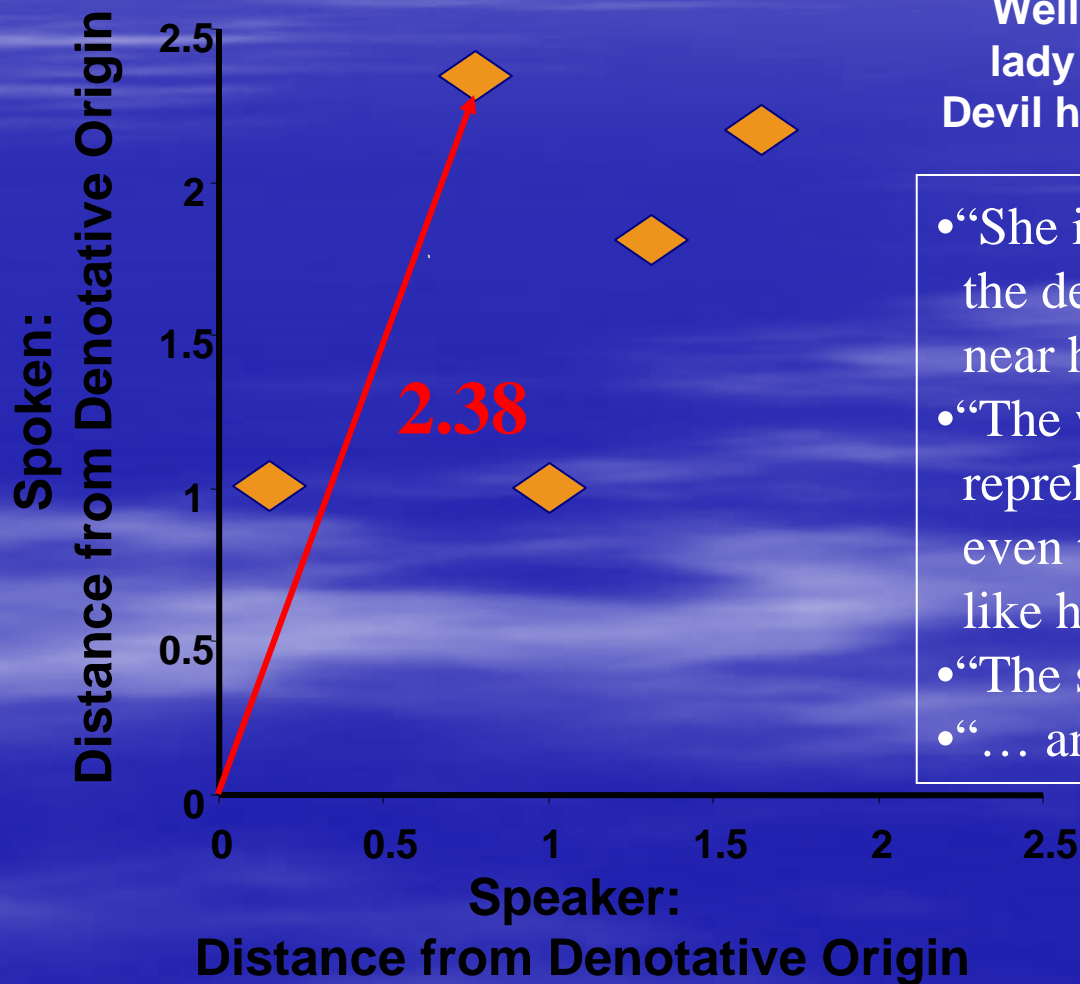
Distance from the Denotative Origin



Well, this wonderful lady would send the Devil himself into hiding.

Degree of Connotation

Distance from the Denotative Origin

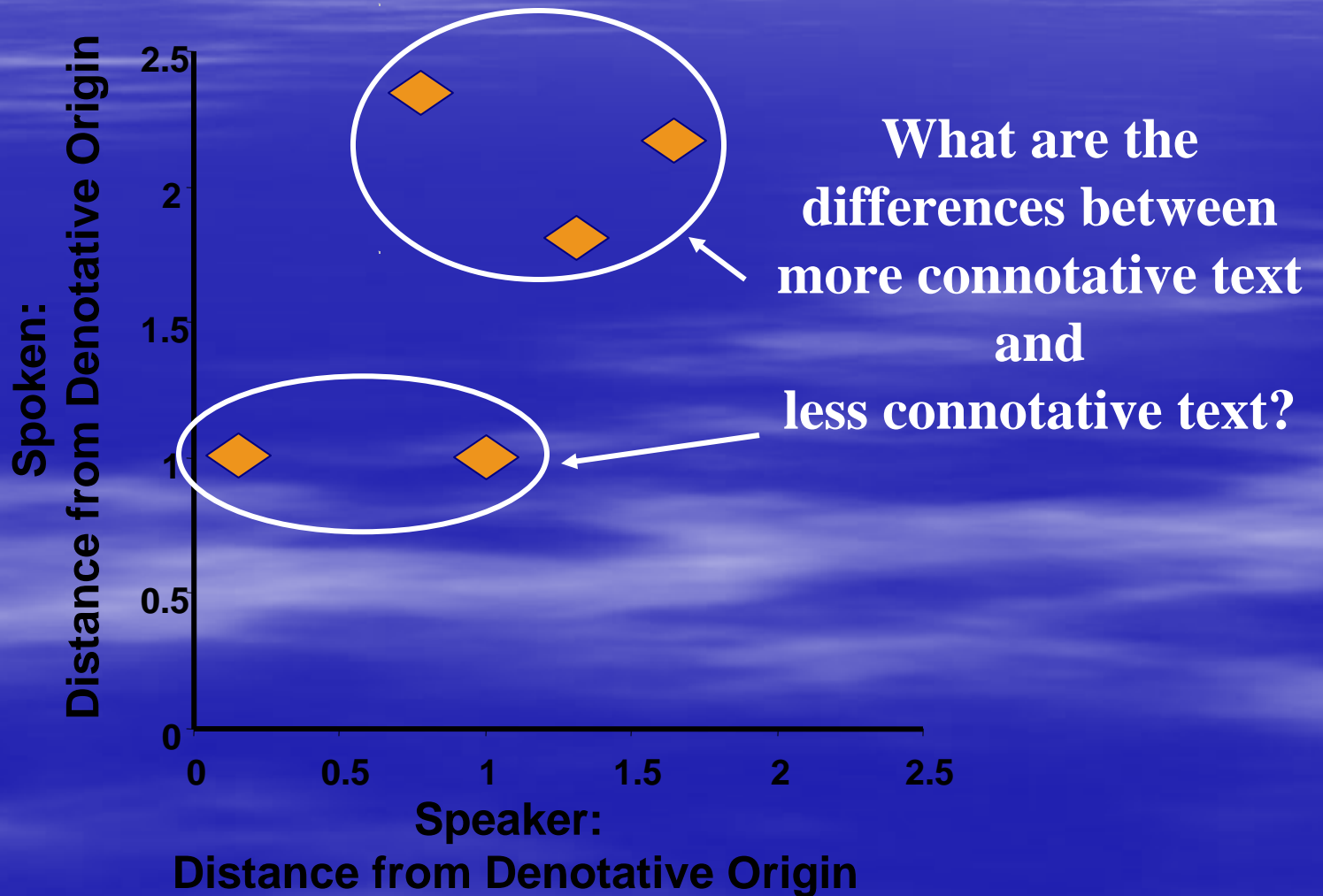


Well, this wonderful lady would send the Devil himself into hiding.

- “She is so good, that the devil wouldn’t go near her.”
- “The woman is so reprehensible that even the devil wouldn’t like her company.”
- “The speaker is older”
- “... and perhaps Irish”

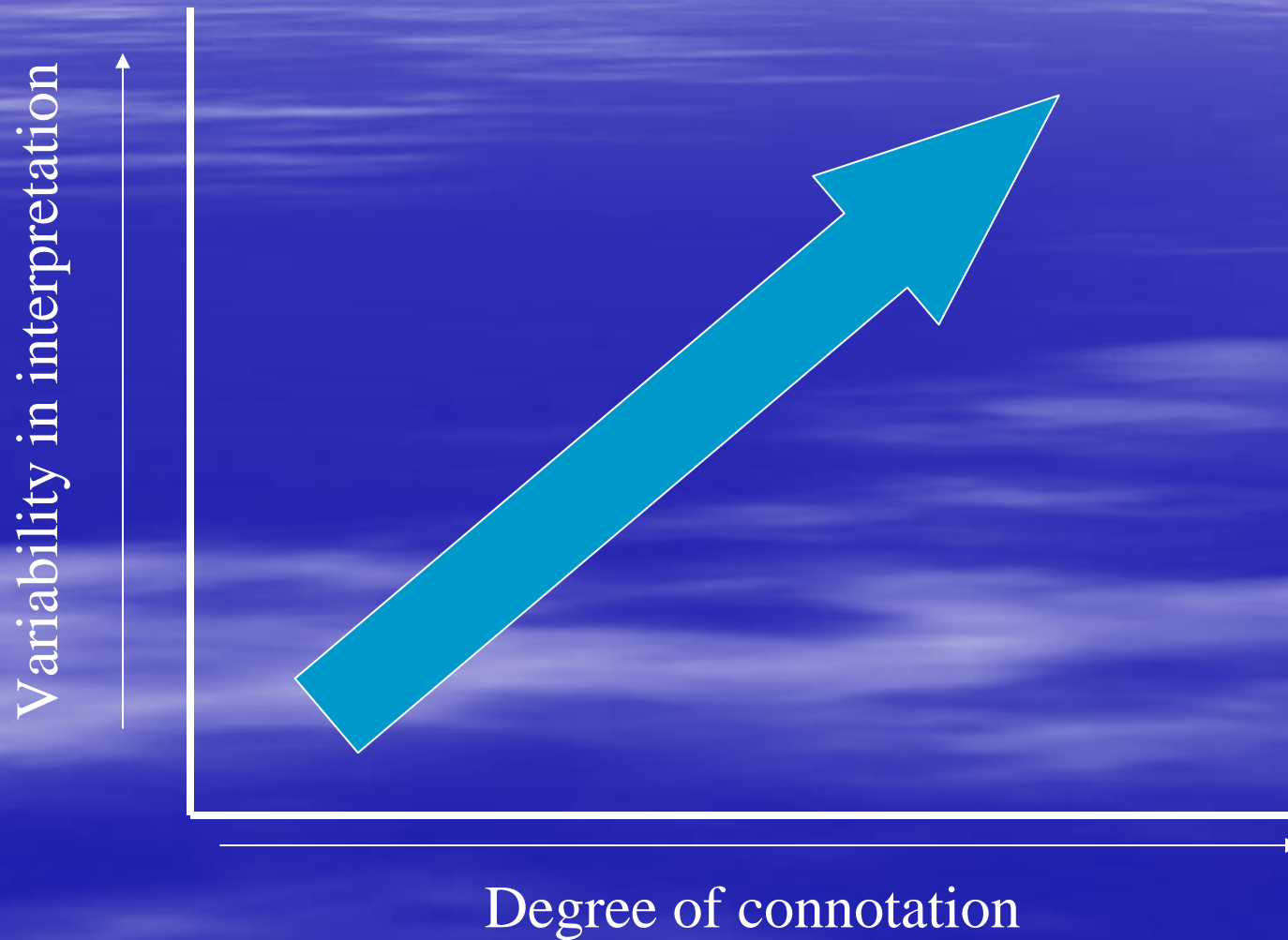
Degree of Connotation

Distance from the Denotative Origin



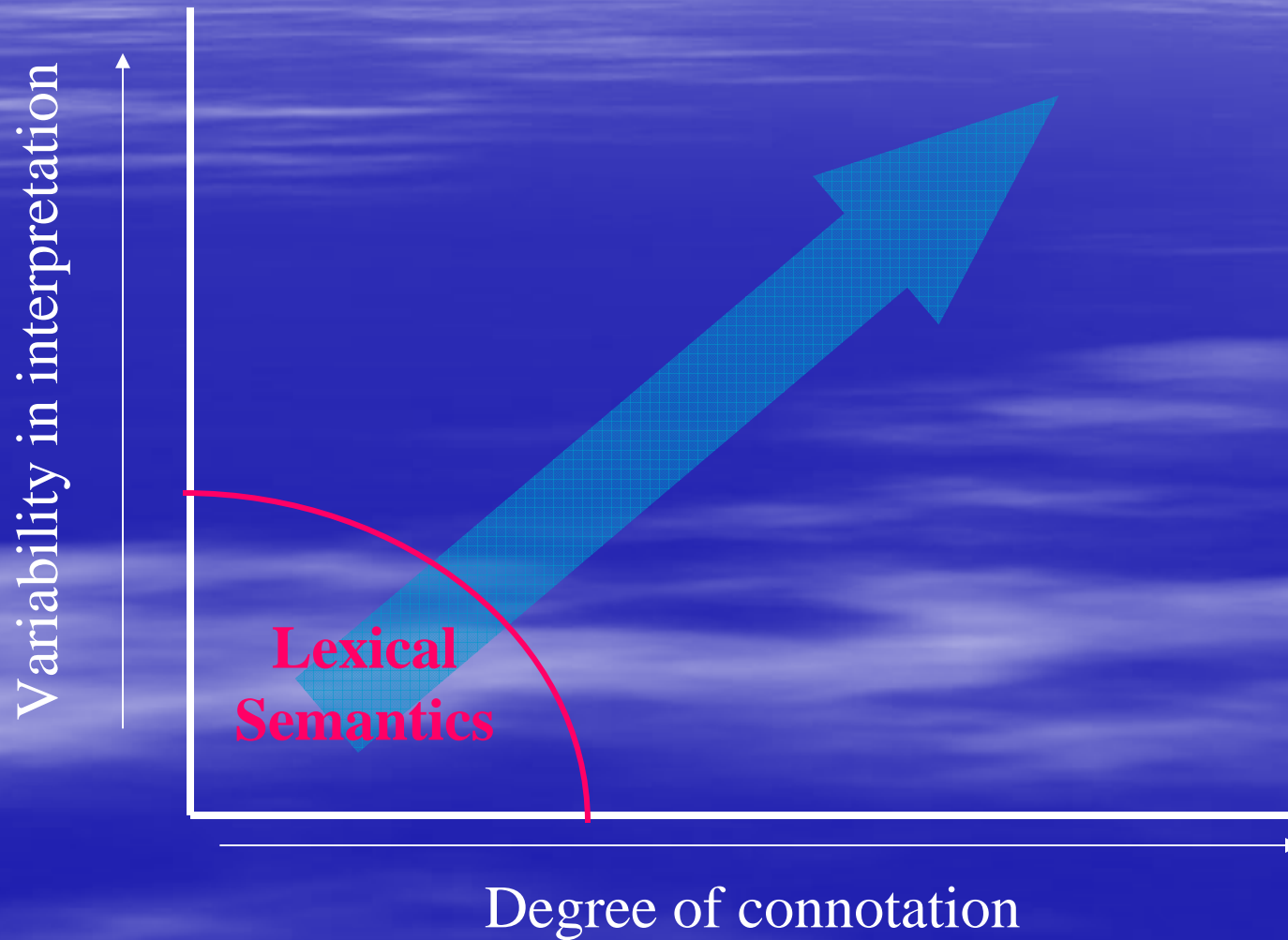
Degree of Connotation

A notion derived from data



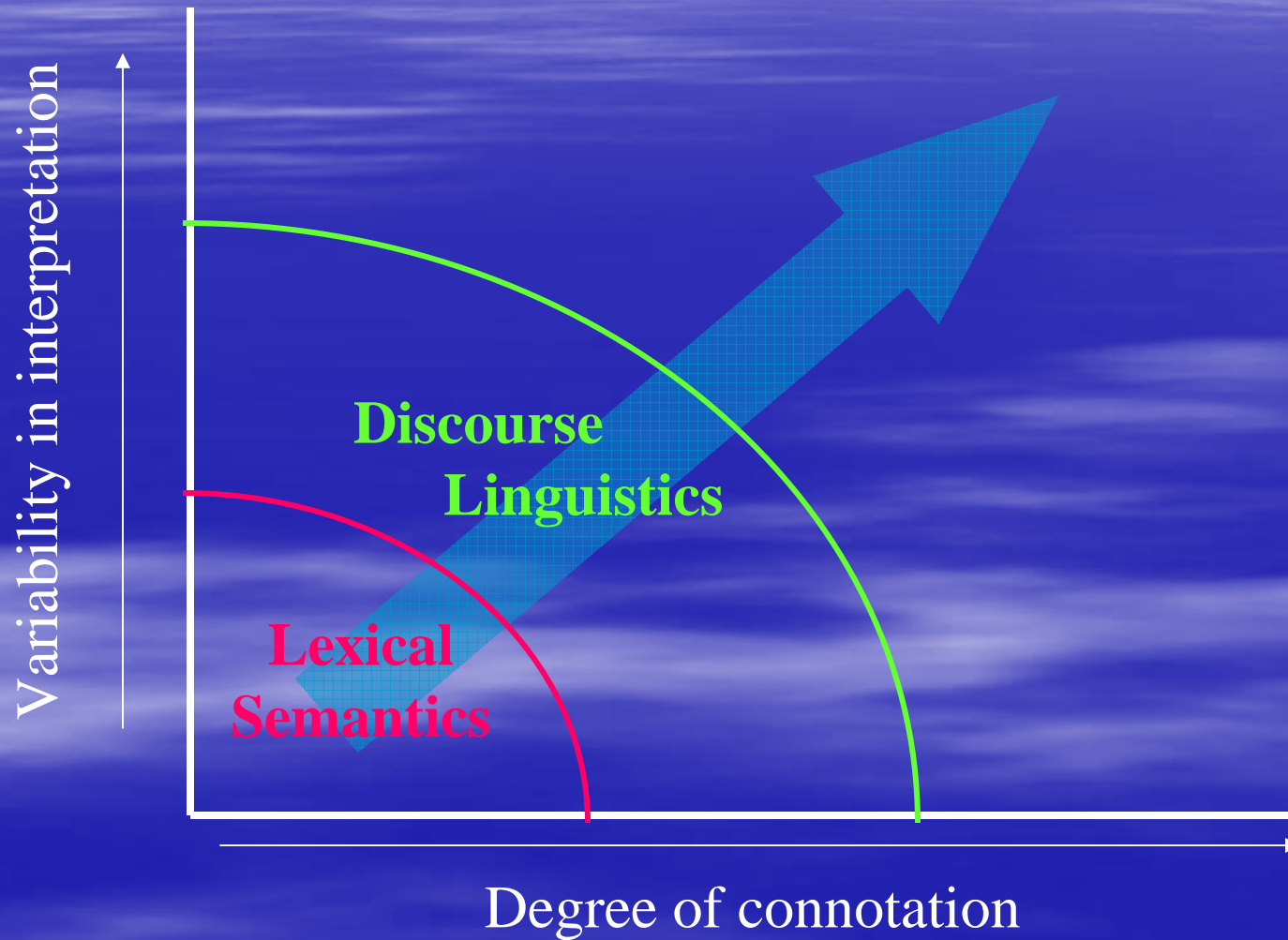
Degree of Connotation

A notion derived from data



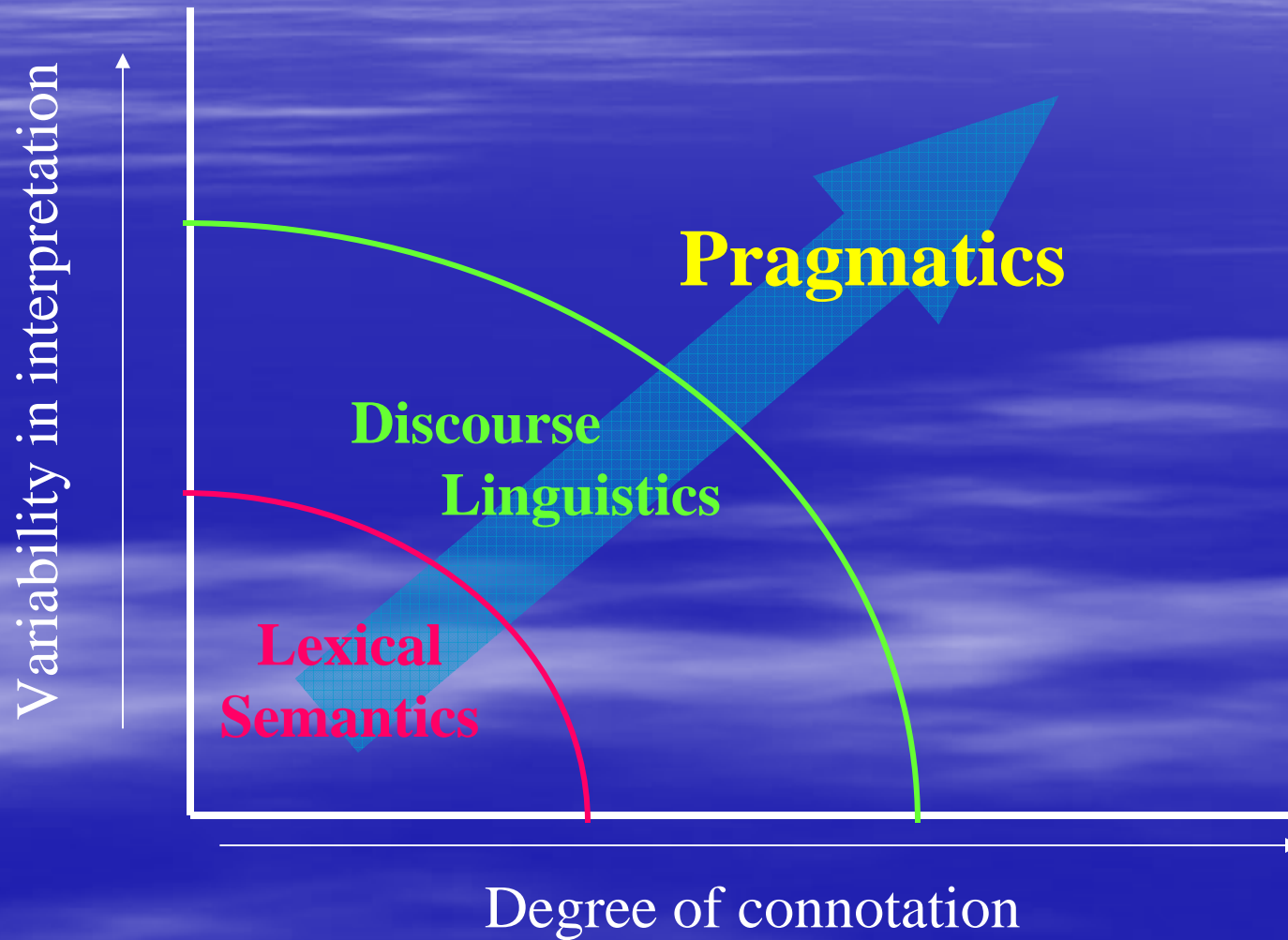
Degree of Connotation

A notion derived from data



Degree of Connotation

A notion derived from data



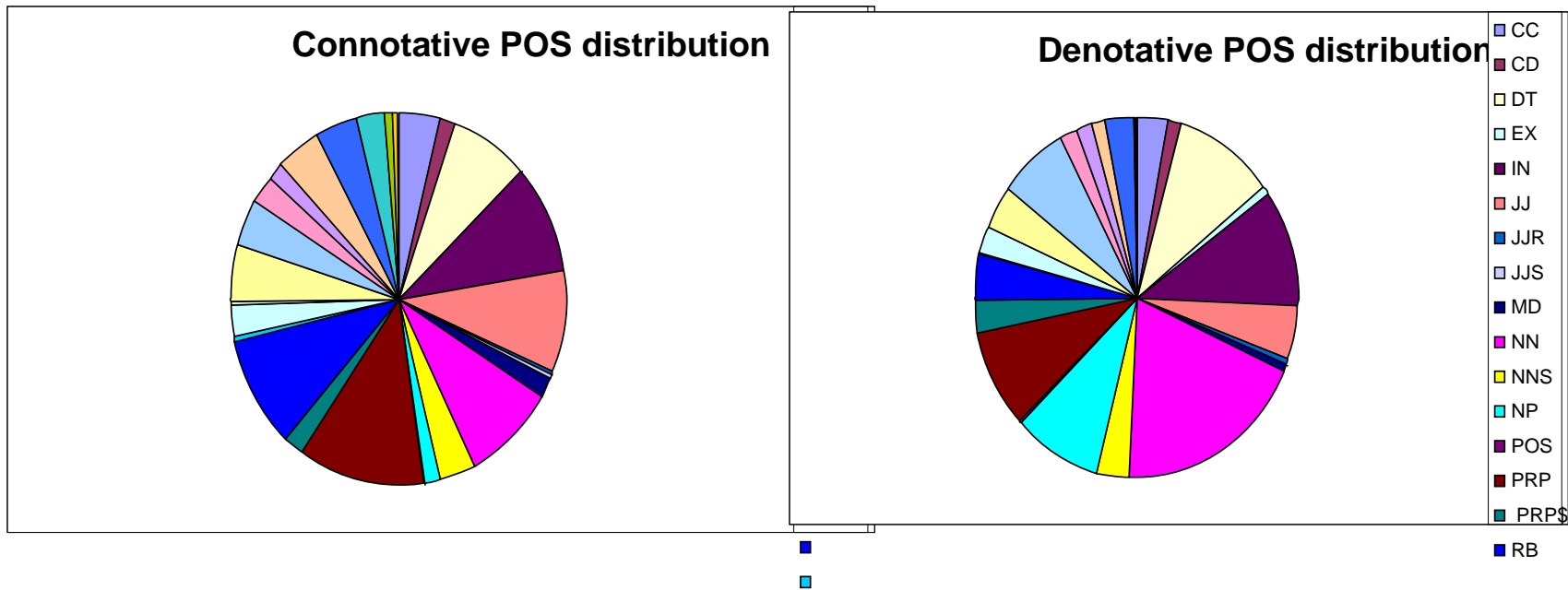
Phase II – TextTagger Feature Identification & Machine Learning Experiments

- All texts judged by subjects will be run through CNLP's TextTagger System
- Will produce feature vector representations using all available levels of language processing
- Linguistic Features
 - POS, tense, person, entities, relations, events
 - Counts of adjectives, adverbs, modals, intensifiers
 - Sentence length / complexity
 - Semantic classes of words & phrases
 - Term collocations
 - Discourse structure
 - Rhetorical structure
- Non-Linguistic Feature
 - Average connotation score across multiple judges

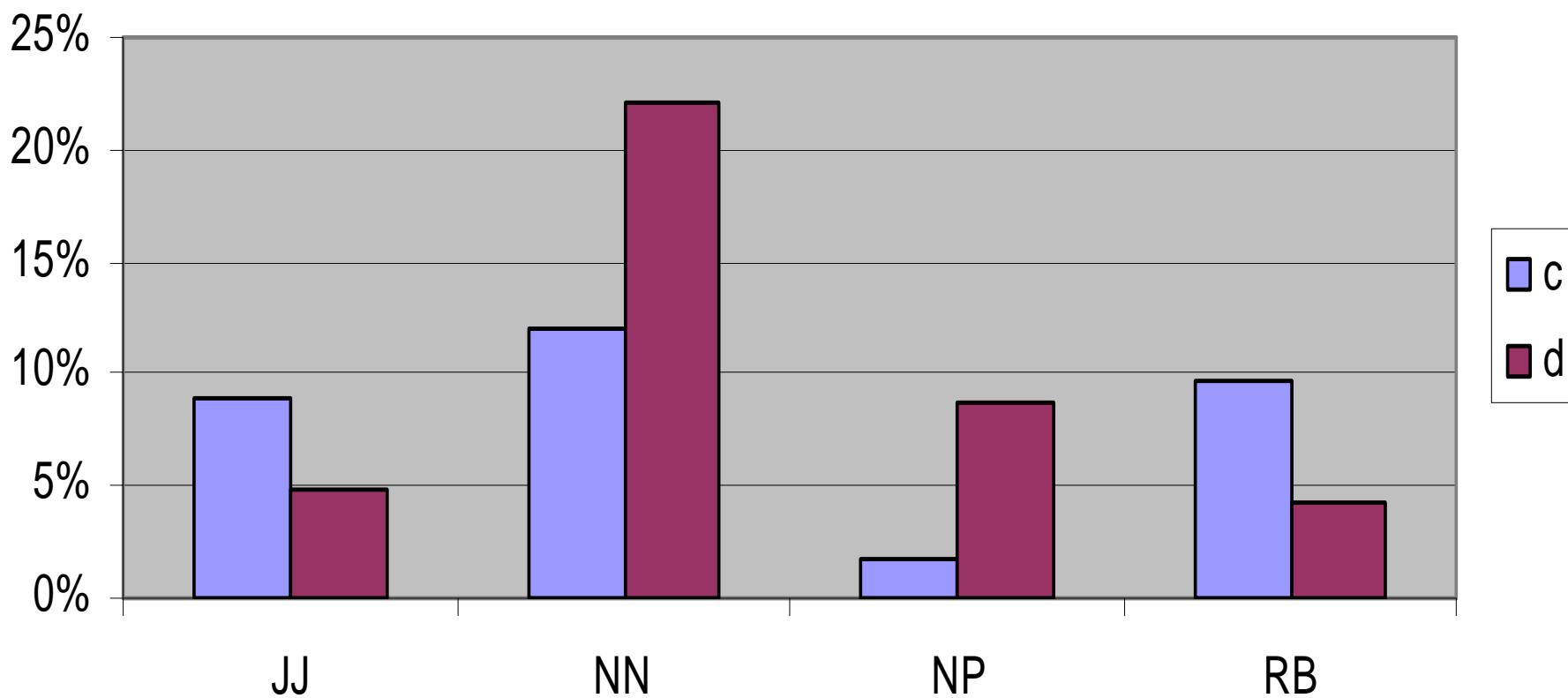
Preliminary Runs on Sample Data

- 10 sentences were provided to 4 subjects who were asked:
 - *What does this sentence suggest?*
 - *What makes you think this?*
- Their responses were coded by 3 researchers as to whether subject saw the sentence as connotative:
 - *"Seems like..."*
 - *"Implies that maybe..."*
- Or denotative:
 - *"Because the sentence says..."*
 - *"Directly in the sentence he refers to..."*
- Sentences run through TextTagger
 - With Connotative / Denotative label as feature

Comparative POS Distribution



POS differences



Characteristics of Connotative Sentences

- Slang
 - *hitting on; he went on & on*
- Idioms
 - *send the Devil himself into hiding*
- Quotation marks & interesting punctuation
 - “gorgeous girlfriend”; !!!
- Hyperbolic Language
 - *what would the two ever possibly have to talk about*
- Hedges, couching
 - *Well; If I had like*

EMPIRICAL SELECTION METHODOLOGY

- Framework for managing many categorization experiments on a variety of dimensions with tools for assessing and analyzing the results
- Enables easy creation of various representations from available NLP tags
- Will use the full set of features both linguistic and non-linguistic, to develop a baseline connotation detector and then use ESM to select the most relevant features for the connotation detection task
- Procedures:
 - Define document representations
 - Create collections for each representation type
 - Use 10-fold cross validation
 - Run paired t-test on results
 - Select best representations based on t-tests

Phase III - Evaluation

- Intrinsic
 - Precision, recall and F_1 -measure for recognition of connotation as compared to human judgments

- Extrinsic
 - We are conducting our work in such a way that we will be able to contribute connotation recognition to QA systems
 - We provide all our technologies as JAVA libraries with fully documented APIs
 - Have experience with UIMA, Web Services & J2EE for integrating with other systems

 - Strong interest in demonstrating how the recognition of connotative meaning can contribute to richer / more subtle question-answering capabilities

Demographics

Total	193
-------	-----

Female	122
Male	71

Education	
Undergraduate	10
Masters	153
PhD	21
Other	9

Native English Speakers	159
-------------------------	-----

Age	
15-20	0
21-30	77
31-40	45
41-50	45
51+	26

Two sets of 10 sentences (split into two presentation orders)	83
	110