

# The Power Beneath

---

*Judith A. Markowitz, PhD*

J. Markowitz, Consultants  
Chicago, IL  
[www.jmarkowitz.com](http://www.jmarkowitz.com)

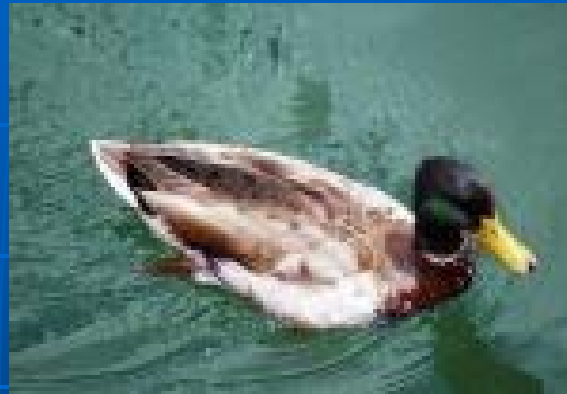
---

AVIOS Voice Search Conference

March, 2008

# On the Surface

---



**The goal/function of the application  
or product**

# The Power Beneath



**The tools and technologies that  
make applications/products more  
accurate and powerful**

# Challenges

---

## NIST TREC-9 (2000)

- Dialogs
- Non-lexical information
- Multi-media

# The Power Beneath

---

## Focus

- What is being communicated (lexical)
- Speaker

# What Is Being Communicated



## On the Surface

- Word spotting
- Indexing and searching the index
- Word-level categorization
- Multi-media search

# What Is Being Communicated



## The Power Beneath

- Dialog management
- Emotion
- Noise management

# What Is Being Communicated



## The Power Beneath

- Clustering and Collocation
- Word Confusion Networks



# What Is Being Communicated



## The Power Beneath

- Clustering
- Collocation
- Affinity

# Clustering & Collocation

**Clustering** grouping of data into sets or groups that share some attributes

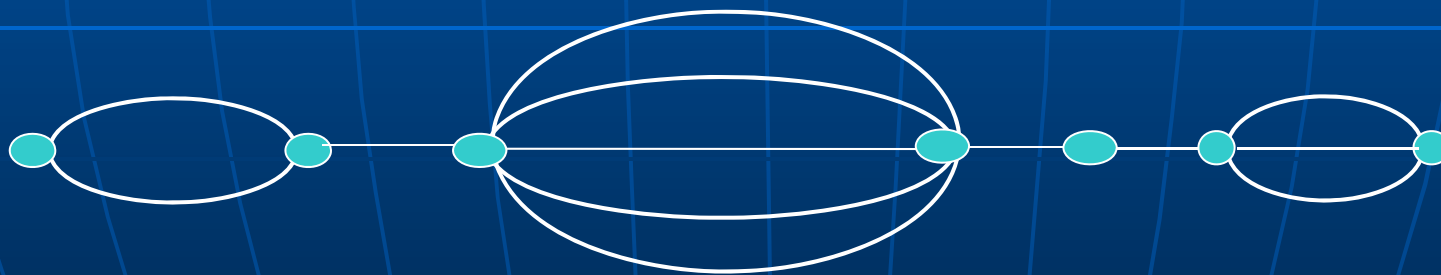
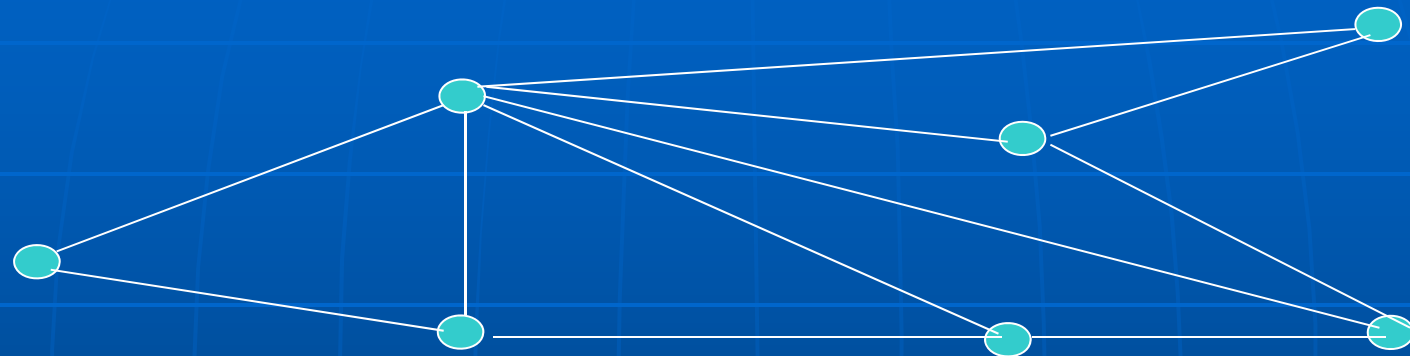
**Collocation** (linguistics) words/phrases that often co-occur. Not necessarily contiguous.

Evens, Litowitz, Markowitz, Smith & Werner 1980 *Lexical-Semantic Relations* Carbondale, IL: Linguistic Research Inc.

**Affinity** (IR) method of ranking content to determine how closely two documents are related.

Zhang, Li, Liu, Ji, Xi, Fan, Chen, Ma 2005 Improving Web Search Results Using Affinity Graphs. In *SIGIR 2005* pp.504-511.

# Word Confusion Networks



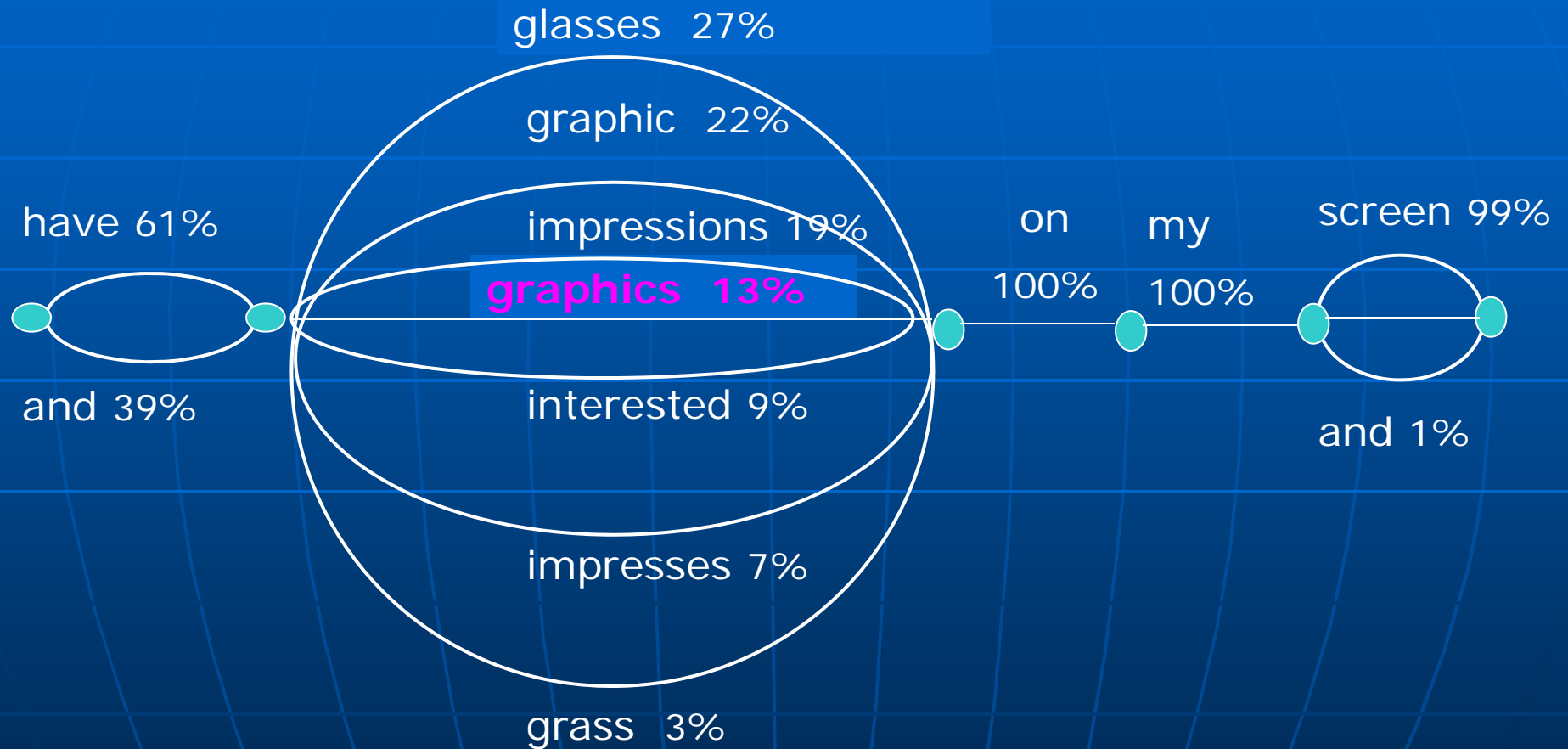
# Word Confusion Networks

---

... the AT&T net client HAVE GLASSES  
ON MY SCREEN fast so it's  
connecting...

From: Marnou, Carmel, Hoory (2006) Spoken Document  
Retrieval from Call-Center Conversations. In SIGIR 2006  
p51-58

# Word Confusion Networks



# Speaker



## On the Surface

- Speaker identification task for security/surveillance/intelligence/law enforcement
- Speaker tracking in a multi-speaker group
- Small set speaker verification/identification for security/privacy

# Speaker



## The Power Beneath

- Black list
- Personalization
- Sales/service
- Classification

---

# QUESTIONS?



# Thank you

---

*Judith A. Markowitz, PhD*

J. Markowitz, Consultants

Chicago, IL

[www.jmarkowitz.com](http://www.jmarkowitz.com)

[Judith@JMarkowitz.com](mailto:Judith@JMarkowitz.com)

+1-773-769-9243