

# **The Uniqueness of Initial Words in News Headlines**

TR 29.3412  
March 2, 2001

Melanie D. Polkosky  
Wallace J. Sadowski  
James R. Lewis

IBM Voice Systems

West Palm Beach, Florida



## **Abstract**

We examined 142 news headlines to determine whether use of the first and second word pair is a feasible news story selection strategy in speech recognition applications. No headlines in the sample contained matching first-second word pairs. A majority of the first two words in the headlines were nouns and 96% of the first-second word pairs included at least one noun. These findings suggest that speaking the first and second word in a headline is a potentially viable news story selection strategy, allowing a user to uniquely select a story of interest in speech recognition applications.

## **ITIRC Keywords**

Voice portals  
Speech recognition application  
News feature  
Headlines  
Selection strategy



# Contents

<b>INTRODUCTION</b> .....	<b>1</b>
<b>METHOD</b> .....	<b>1</b>
<b>RESULTS</b> .....	<b>3</b>
<b>DISCUSSION</b> .....	<b>5</b>
<b>REFERENCES</b> .....	<b>7</b>
<b>APPENDIX A. HEADLINE SAMPLES (FIRST-SECOND WORD PAIRS)</b> .....	<b>9</b>
<b>APPENDIX B. HEADLINE SAMPLES (FULL TEXT)</b> .....	<b>13</b>



## **Introduction**

News reports are a common feature of voice portal information services (Manes, 2000; Mullaney, 2000). Currently, voice portals (including BeVocal<sup>1</sup>, Tellme<sup>2</sup>, and HeyAnita<sup>3</sup>) allow users to select from 6 to 8 categories of news and listen to several news stories in a pre-determined sequence. These speech recognition applications allow minimal navigation between news stories and do not provide headlines to announce the stories.

Chang (1998) found that news teasers (a description or preview of an upcoming news item) enhance viewer comprehension and recall of television news stories. In an information service that uses voice recognition, headlines may function as teasers, aiding users' comprehension and recall of news stories. They may allow users to rapidly select specific news stories of personal interest when words from a headline are spoken to the speech interface. However, no studies have investigated different headline selection strategies for speech recognition applications.

In this study, we examined one potential technique for news story selection using headlines. Specifically, the purpose of this study was to determine the feasibility of using the first one to two words of a headline to select a desired news story without interference from similar headlines.

## **Method**

We obtained 142 headlines from a major Internet news provider for news stories (<http://www.cnn.com/>). The news stories appeared on 3 separate days during February 2001. We chose the headlines from categories similar to those offered by voice portal news features: Top Stories, Business, Sports, News by Region, and Science and Technology. We examined the headlines to determine whether the same words occurred in the first or second position in the headlines or whether both the first and second words matched. Analysis consisted of counting the number of matching single words (first words, second words) and first-second word pairs. An additional analysis determined the grammatical part of speech of the first-second word pairs.

---

<sup>1</sup> BeVocal is a trademark of BeVocal, Inc.

<sup>2</sup> Tellme is a registered trademark of Tellme Networks, Inc.

<sup>3</sup> HeyAnita is a trademark of HeyAnita, Inc.





## Results

Table 1 shows the number of first word, second word, and first-second word pair matches in the headline sample.

*Table 1. Matching Single Words and First and Second Word Pairs*

	<b>1st Headline Word</b>	<b>2nd Headline Word</b>	<b>1st &amp; 2nd Headline Words</b>
<b>Matching Pairs</b>	13	2	0
<b>Matching Percentage (90% CI)</b>	9.15 (5.5-14.2)	1.41	0 (0-2.1)

We found matching first words 13 times (9.15% of the time) and no matching first and second word pairs. Of the matching single first words, “U.S.” occurred in 4 headlines and “report,” “Pentagon,” and “NTSB” each occurred in 3 headlines. Two headlines included the single words “AOL,” “demonstrators,” “latest,” and “new.”

Table 2 shows the grammatical part of speech of each of the first two words in the headline sample.

*Table 2. Grammatical Classes of First-Second Word Pairs*

<b>Part of Speech</b>	<b>Noun</b>	<b>Verb</b>	<b>Adjective</b>	<b>Adverb</b>	<b>Preposition</b>	<b>Conjunction</b>
<b>Number of Occurrences</b>	194	49	20	7	13	1
<b>Percent of Total Words</b>	69.3	17.3	7.0	2.5	4.6	0.4

We coded the part of speech based on the lexical characteristics of each word, not its function in the headline; for example, a proper noun functioning as an adjective (e.g., “Army report”) was coded as a noun. We found that most first and second words are nouns (69%), with verbs occurring 17% of the time, and other grammatical classes minimally represented. In addition, 136 of the 142 first-second word pairs (96%) included at least one noun.



## **Discussion**

This study investigated the feasibility of using first and second word pairs from headlines as a selection strategy for news stories in speech recognition applications. An examination of 142 headlines sampled from a major Internet news service showed that no headlines contained the same first and second word pair. This finding indicates that the first-second word pair uniquely represents a single news story and will allow accurate retrieval of that story in speech recognition applications.

We also found that 69% of the first two words in the headlines are nouns and 96% of the first-second word pairs include at least one noun. Simpson & Klippert (1968) found that nouns are most easily learned of all grammatical classes in lists, especially when the lists have sentence word order. Therefore, a second potential benefit of the first-second word selection strategy is ease of recall and use by users.

Because the current study uses headlines from 3 days of news stories, a single days' news headlines would be less likely to match for single words than our findings suggest. In general, this study indicates that speaking a first-second word pair from a headline is a viable selection strategy in speech recognition applications.



## References

Chang, H. (1998). The effect of news teasers in processing TV news. *Journal of Broadcasting and Electronic Media*, 42(3), 327-339.

Manes, S. (2000, July 24). The web speaks, sort of. *Forbes*, 212-214.

Mullaney, T. (2000, Nov. 13). Vortal combat: The early years. *Business Week*, 146-148.

Simpson, W. & Klippert, K. (1968). Grammatical class differences in serial learning. *Journal of General Psychology*, 79(2), 183-189.



## Appendix A. Headline Samples (First-Second Word Pairs)

First Word of Headline	Second Word of Headline
30	die
45	years
Moon	hoax
Analysis	Interference
AOL	looks
AOL	Time
Ariba	stock
Army	report
Asian	stocks
Astronauts	unfold
At	least
AT&T	finishes
Attorneys	for
Best	students
Big	bang
Bill	Gates
Brazil's	Cardoso
Broomstick	faulty
Bush	nominee
Cable	vs.
California	power
Chiefs	officially
Clinton	seeks
Coke	adds
Columbine	High
Commercial	actors
Convicted	terrorist
Crippled	destroyer
Dan	Ronan
Death	row
Defense	in
Defunct	online
Delta	seeking
Demonstrators	mark

Demonstrators	protest
Detroit	apartment
Dozens	arrested
Election	showdown
Endeavour	preparing
EU	unveils
Europe	rallies
Execution	set
FAA	proposes
Fed	cuts
Federal	study
Flight	diverted
Florida	high
For	Missouri-born
Former	tennis
France	weighs
French	sell-off
Gas	prices
Ginger	inventor
GOP	tax-cut
Grocery	chain
Heavy	snow
Hewlett-Packard	co-founder
Hokies	Vick
Hubble	spots
In	Focus
Indiana	church
Internet	Explorer
Iranian	hanged
Israeli	Palestinian
Jack	Hemingway
Jackson	sounds
Japan	political
Key	Democrats
Latest	attacks
Latest	earnings
La-Z-Boy	Microsoft
Legend	Airlines
Leonard	races



Lewinsky	will
Lockerbie	victims
Loverboy	bass
Man,	woman
Mars	Global
Matter	seen
Miami	police
Mideast	violence
Millions	of
Minor	planet
Mobile	Playstation
MSFT	appeal
Much	hyped
NASA	Satellite
NATO	to
NBA	fines
NBC	to
New	CEO
New	York
News	from
No	injuries
Nokia	sees
NTSB	chairman
NTSB	changes
NTSB	concerned
Oakland	concerned
Officer	gets
Oklahoma	farmer
Osama	bin
Pentagon	says
Pentagon	spokesman
Pentagon	to
Philip	Morris
Pinochet	enters
Prosecution	could
Prosecutors	finish
Pulsar	linked
Pumps	to
Quake	shakes

Report	blames
Report	Pitino
Report	Too
Robbery	suspect
Salt	Lake
Sandia	labs
Scooter	wave
Scout	leader
See	you
Sega	dumps
Senate	debating
Silicon	Valley
South	African
Spinal	surgeon
Stamp	prices
Staples	installing
Stocks	slide
Stoops	interested
Talks	to
Tanker	to
Toll	estimates
Treasurys	still
U.S.	Army
U.S.	could
U.S.	growth
U.S.	launches
Vikings	and
Wall	St.
Web	based
Where	did

## Appendix B. Headline Samples (Full Text)

Wall St. pulls back  
U.S. growth slows  
Latest earnings news  
French sell-off tripped up  
Treasury's still rising  
In Focus: AOL  
New CEO at Chrysler's wheel  
News from Latin America  
Philip Morris falls short  
Bill Gates on world health  
Stocks slide on Wall St.  
Europe rallies behind techs  
Asian stocks end higher  
AOL looks ahead, but stock drops  
Coke adds bottled coffee  
AT&T finishes Excite swap  
MSFT appeal challenged  
Ariba stock takes a dive  
Hewlett-Packard co-founder dies  
NBC to pare work force  
Fed cuts interest rates  
Lockerbie victims' relatives want Gadhafi punished  
U.S. could be stuck with California utilities' debts, senator warns  
Toll estimates rise as quake rescuers depart  
Talks to tackle N. Ireland impasse  
Broomstick, faulty electric fence aid Alabama jail break  
Mideast violence underscores final days of Israeli campaign  
GOP tax-cut push buoyed by growing surplus estimates  
Senate debating Ashcroft nomination  
Key Democrats challenge Bush's education plan  
Election showdown today at U.S. Supreme Court  
Clinton seeks to increase child immunizations  
Report blames Firestone  
Man, woman killed in crash of light aircraft in Oklahoma  
NASA: Satellite odds of hitting someone 250-1  
Best students anxiously await word on early acceptance to college  
California power crisis eases a bit, but still critical  
AOL Time Warner goes to work  
Silicon Valley pioneer William Hewlett dead at 87  
Bush nominee for secretary of interior bombarded by critics  
Cable vs. DSL  
Mobile Playstation  
Sega dumps Dreamcast game machine

Mars Global Surveyor completes mapping mission  
U.S. launches new Global Positioning Satellite  
'Ginger' inventor at Davos, stays mum  
Internet Explorer 6 gets an early outing  
Staples installing Internet kiosks  
Pumps to deliver more than just gas  
EU unveils plan to fight cybercrime  
Matter 'seen' disappearing into black holes for first time  
Web based education on a slow, steady path  
La-Z-Boy, Microsoft launch 'e-cliner' armchair  
Big bang balloon brought back  
Hubble spots 'pipeline' between rowdy galactic neighbors  
Pulsar linked to ancient star explosion recorded by Chinese  
'Moon hoax' spurs crusade against bad astronomy  
Tanker to dock with Mir before space station's demise  
Analysis: Interference issues hinder Bluetooth  
Much hyped 'Ginger' may be motorized scooter  
Report: Pitino, UNLV close to finalizing deal  
Chiefs officially name Vermeil head coach  
Stoops interested in Browns' vacancy  
Leonard races to Mercedes lead; Woods 3 back  
Jackson sounds off on Lakers' power struggle  
NBA fines Mavs owner Cuban again  
Hokies' Vick finally makes NFL plans official  
See you in Tampa  
Vikings and Raiders confident they can prevail  
Oakland concerned about Baltimore's swarming 'D'  
Latest attacks claim 9 lives in violence-torn Algeria  
South African officials fear cholera epidemic might spread  
At least 113 deaths blamed on Asian winter weather  
Millions of saints, sinners, citizens bathe in Ganges at Hindu festival  
Japan political woes loom amid faction split fears  
Brazil's Cardoso cuts fine from anti-slander law  
Pinochet enters hospital for mental and neurological tests  
30 die but Colombia rebels said to plan peace move  
NATO to probe uranium threat  
France weighs in over German BSE crisis  
Prosecutors finish closing arguments in Lockerbie bombing trial  
Iranian hanged in public for killing member of militia  
No injuries reported in small blast at Saudi mall  
Osama bin Laden shown at wedding of son  
Nokia sees 1 billion mobile users by mid-2002  
Israeli, Palestinian violence continues in West Bank, Gaza  
Stamp prices going up next month  
Crippled destroyer due in U.S. port next week, Navy chief says  
Lewinsky will 'reluctantly' appear for interview by prosecutors

Miami police: O.J. Simpson involved in altercation with motorist  
Federal study shows kindergarten improves all young minds  
Report: Too few programs to help juveniles with mental health problems  
Florida high court to hear Gore appeal tomorrow  
Astronauts unfold space station's second solar wing  
'Minor planet' found in obscure corner of solar system  
Defense in Lockerbie opens case against murder and terrorism charges  
Delta seeking court order against pilots union  
Flight diverted after man's comment about bomb  
Former tennis champ Boris Becker splits from wife  
Scooter wave slides across America  
Salt Lake organizers reveal elaborate Olympic torch run  
Execution set for tonight would set record in Texas  
Attorneys for convicted Los Angeles officers will seek new trial  
FAA proposes almost \$1 million fine for Alaska Airlines  
Indiana church plans Christmas Eve Mass in shadow of Colt's game  
Sandia labs signs agreement with Navajo Nation  
Heavy snow in U.S. Great Lakes region  
Where did TV's e-tailer ads go?  
NTSB changes opinion on AirTran cooperation  
Pentagon to notify Gulf veterans of possible nerve gas exposure  
Pentagon spokesman Bacon leaving to head world refugee group  
Gas prices take seasonal dip  
For Missouri-born diver, Russian sub mission was like no other  
Oklahoma farmer, world's oldest man, dies at 111  
U.S. Army schedules troop rotations in Bosnia for next 5 years  
Convicted terrorist to face new trial in Jordan  
Commercial actors vote to officially ratify contracts  
Grocery chain expands recall of ground beef amid E. coli scare  
Robbery suspect shot by would-be victim: armed, off-duty deputy  
Spinal surgeon charged with alleged sexual assault of friend  
Loverboy bass player swept overboard, missing  
45 years later, Rosa Parks inspires without words at museum dedication  
Endeavour preparing for rendezvous with international space station  
Officer gets probation for role in fatal raid  
Jack Hemingway, outdoorsman and son of author, dies at 76  
Legend Airlines suspends operations  
Quake shakes northern California; no injuries reported  
Scout leader says New York chapter not anti-gay  
Death row inmate nominated for Nobel Peace Prize  
Columbine High wins second straight state football title  
Demonstrators mark WTO protest with new march through Seattle streets  
Pentagon says it expects Russian bomber missions near Alaska soon  
Dozens arrested at WTO protest in Seattle  
Defunct online retailer throws lifeline to Alaska sled dogs  
NTSB chairman criticizes AirTran, Atlanta airport over jet fire

Prosecution could get USS Cole bombing case this weekend  
Demonstrators protest new IMF visitor center on Seattle anniversary  
NTSB concerned evidence was washed away after AirTran fire  
Detroit apartment fire kills 6 children  
Army report: Favorable treatment for soldiers accused of misconduct in Kosovo  
Dan Ronan on the investigation into AirTran fire  
New York schools cut ties with Boy Scouts