

Merano, Italy  
December 13-17, 2009



2009 IEEE

AUTOMATIC SPEECH RECOGNITION  
AND UNDERSTANDING WORKSHOP

# Progress and Prospects for Speech Technology *(results of a survey)*

**Roger K. Moore**

*Chair of Spoken Language Processing*  
Dept. Computer Science  
University of Sheffield, UK



ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 1



## The Survey(s)

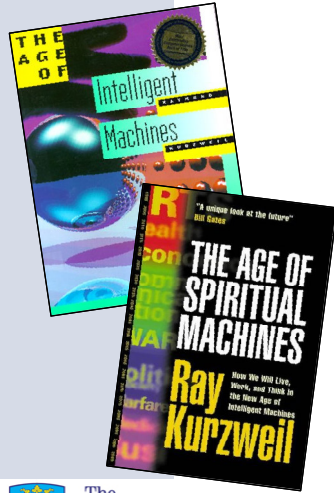
- This is the third 'sexennial' survey to take place at an ASR workshop (*previous surveys were conducted at ASRU-1997 and ASRU-2003*)
- The survey(s) consist of a number of 'statements', such as ...
  - "More than 50% of new PCs have dictation on them"
  - "No more need for speech research"
- Participants are invited to suggest the year (or "never") in which each statement might become true
- The statements have been the same each time, but the number has grown: 12 → 20 → 26
- The statements were suggested by the three ASRU organising committees + some were taken from Ray Kurzweil



ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 2



## Ray Kurzweil's Predictions



**“2009:** *Most routine business transactions take place between a human and a virtual personality (including an animated visual presence that looks like a human face).”*



The University Of Sheffield.

ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 3



## Thanks to All 127 Participants



*“It’s hard to predict ...”*  
*“... especially the future!”*

**Niels Bohr (1922)**



The University Of Sheffield.

ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 4



# The Respondents

## Attending ASRU 2009?

Yes	31	20%
No	92	60%
No Answer	31	20%

## ASRU 2003 Survey?

Yes	10	6%
No	115	75%
No Answer	29	19%

## ASRU 1997 Survey?

Yes	6	4%
No	117	76%
No Answer	31	20%

# The Respondents

IEEE	69	
ISCA	105	
National	31	
Other	29	

# The Respondents

University	100	
Large company (>100 employees)	45	
Small company (<100 employees)	25	
Research institute	29	
Government	6	
Self-employed	4	
Unemployed	2	
Other	0	



The University Of Sheffield.



# The Respondents

Researcher	109	
Developer	34	
Manager	22	
Student	44	
Consultant	11	
Other	3	



The University Of Sheffield.



## Some Overall Statistics

	2009	2003	1997
<b>Respondents:</b>			81
<b>Overall Median:</b> <i>(1<sup>st</sup> 12 statements)</i>			2010
<b>“Never”s:</b> <i>(1<sup>st</sup> 12 statements)</i>			17%
<b>Named responses:</b>			22%



## Some Overall Statistics

	2009	2003	1997
<b>Respondents:</b>		105	81
<b>Overall Median:</b> <i>(1<sup>st</sup> 12 statements)</i>		2020	2010
<b>“Never”s:</b> <i>(1<sup>st</sup> 12 statements)</i>		22%	17%
<b>Named responses:</b>		4%	22%



## Some Overall Statistics

	2009	2003	1997
<b>Respondents:</b>	<b>127</b>	105	81
<b>Overall Median:</b> <i>(1<sup>st</sup> 12 statements)</i>		2020	2010
<b>“Never”s:</b> <i>(1<sup>st</sup> 12 statements)</i>		22%	17%
<b>Named responses:</b>		4%	22%



## Some Overall Statistics

	2009	2003	1997
<b>Respondents:</b>	<b>127</b>	105	81
<b>Overall Median:</b> <i>(1<sup>st</sup> 12 statements)</i>	<b>2028</b>	2020	2010
<b>“Never”s:</b> <i>(1<sup>st</sup> 12 statements)</i>		22%	17%
<b>Named responses:</b>		4%	22%



## Some Overall Statistics

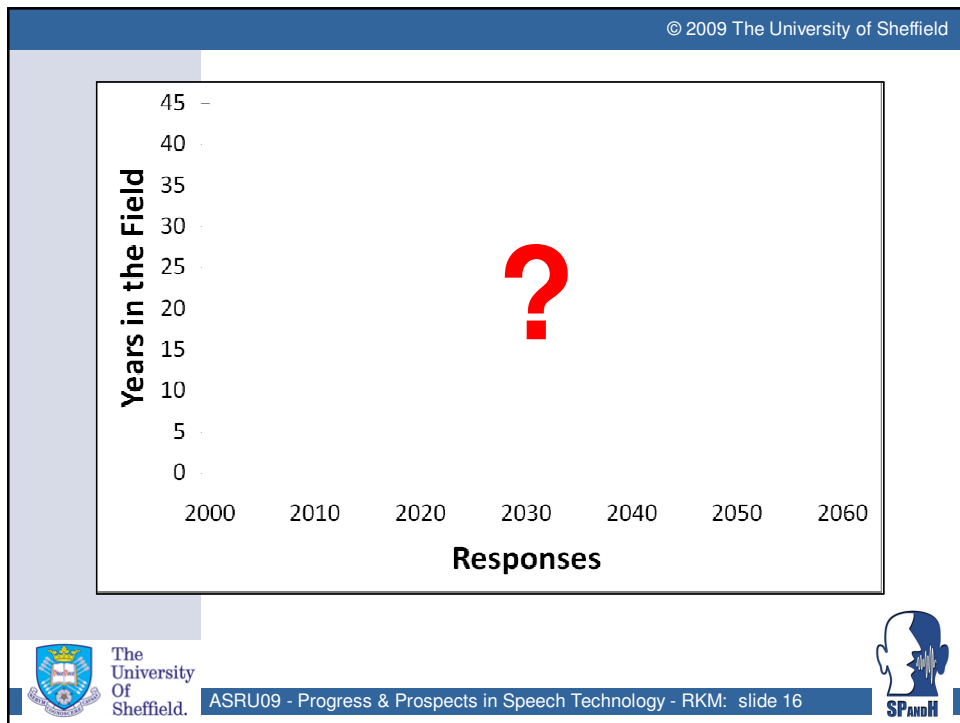
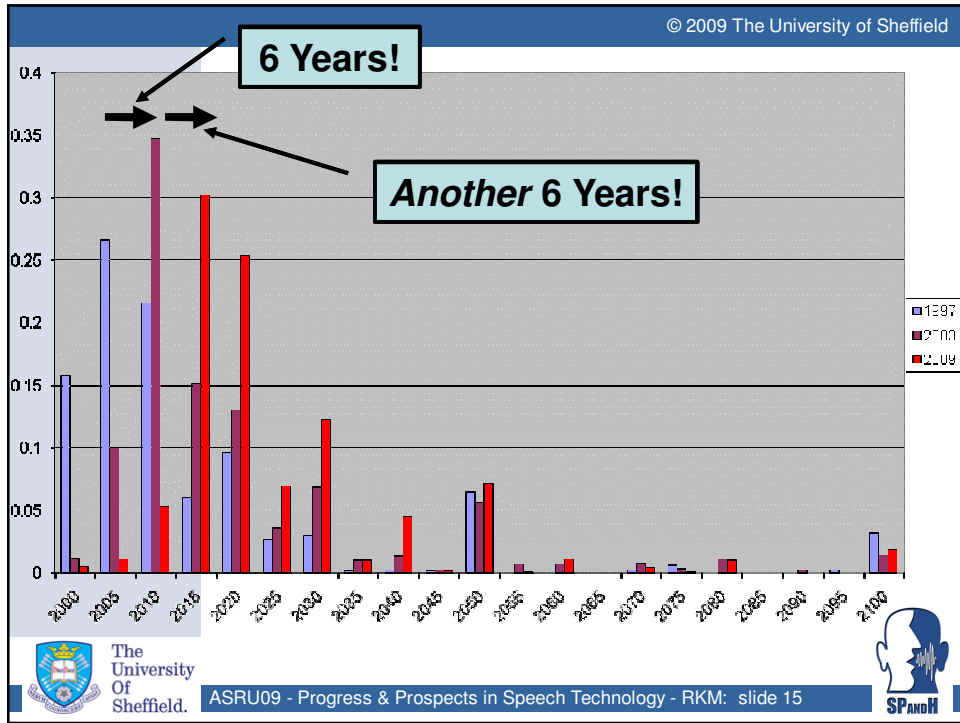
	2009	2003	1997
<b>Respondents:</b>	<b>127</b>	105	81
<b>Overall Median:</b> <i>(1<sup>st</sup> 12 statements)</i>	<b>2028</b>	2020	2010
<b>“Never”s:</b> <i>(1<sup>st</sup> 12 statements)</i>	<b>22%</b>	22%	17%
<b>Named responses:</b>		4%	22%



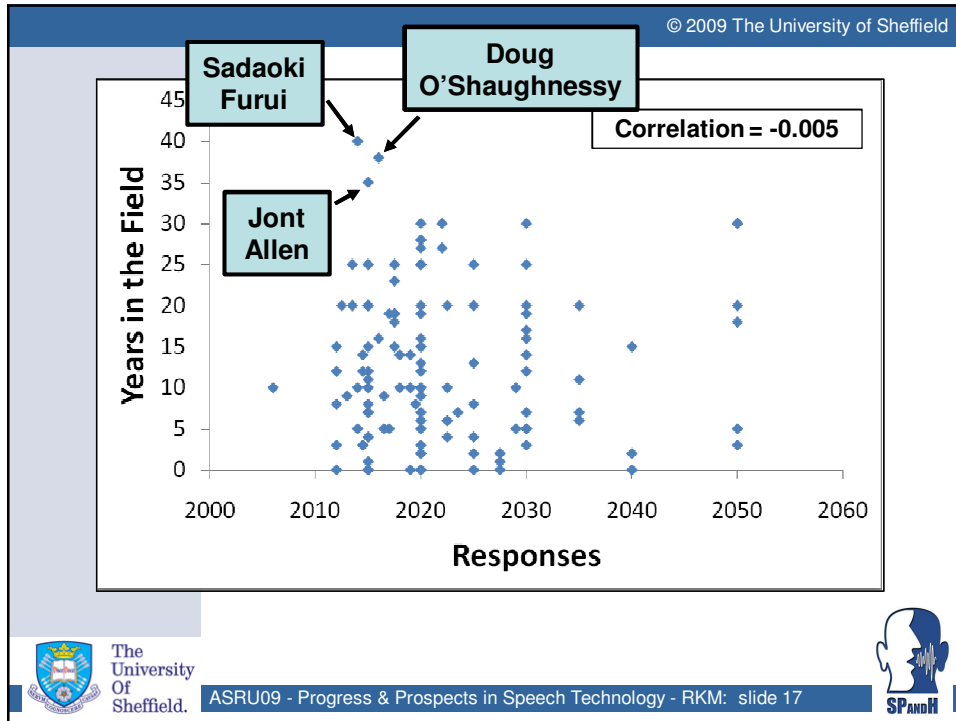
## Some Overall Statistics

	2009	2003	1997
<b>Respondents:</b>	<b>127</b>	105	81
<b>Overall Median:</b> <i>(1<sup>st</sup> 12 statements)</i>	<b>2028</b>	2020	2010
<b>“Never”s:</b> <i>(1<sup>st</sup> 12 statements)</i>	<b>22%</b>	22%	17%
<b>Named responses:</b>	<b>21%</b>	4%	22%









© 2009 The University of Sheffield

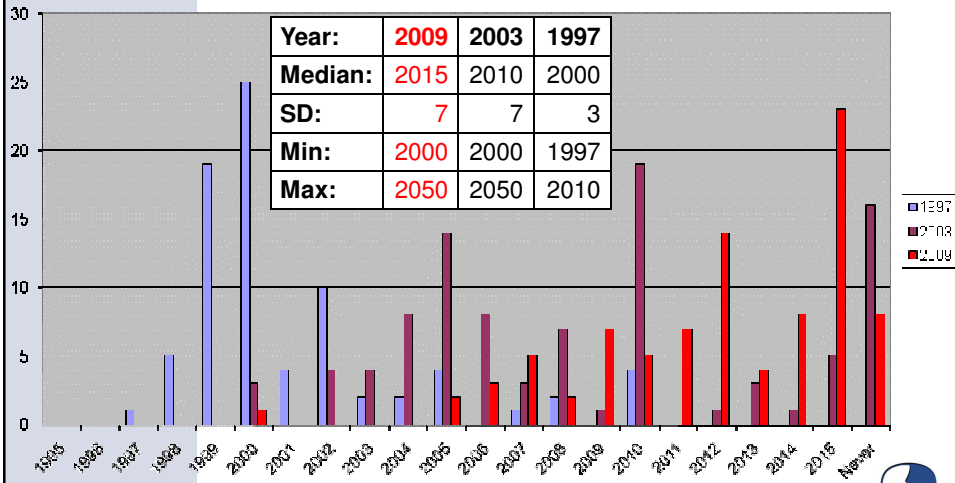
More than 50% of new PCs have dictation on them, either at purchase or shortly after

***“already comes with Vista/Win7”***

***“The real question is when will +50% of people use it?”***

The University of Sheffield. ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 18

**1. More than 50% of new PCs have dictation on them, either at purchase or shortly after.**

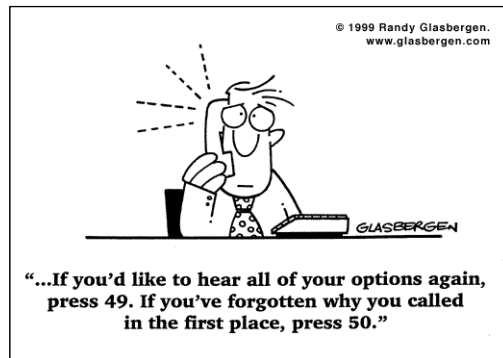


The University of Sheffield.

ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 19



**Most telephone Interactive Voice Response (IVR) systems accept speech input**

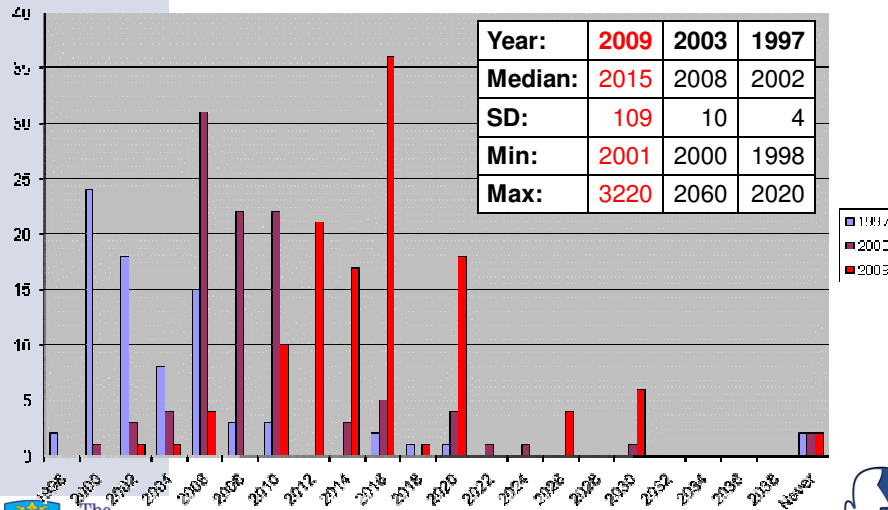


The University of Sheffield.

ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 20



**2. Most telephone Interactive Voice Response systems accept speech input (and more than just digits)**



The University Of Sheffield.

ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 21



Automatic airline reservation by voice over the telephone is the norm

***“via browser over the internet without speech”***

***“web is better”***

***“internet will take over”***

***“phone system will need to be more of a travel agent”***

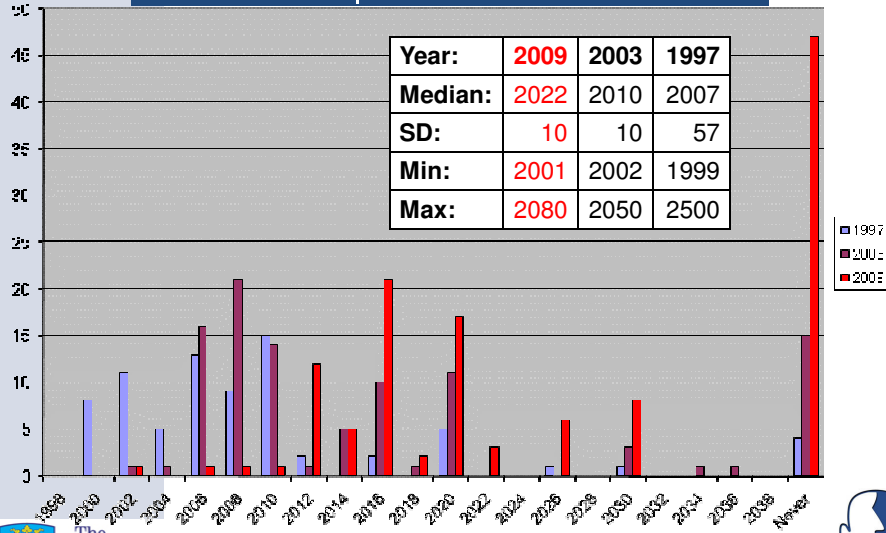


The University Of Sheffield.

ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 22



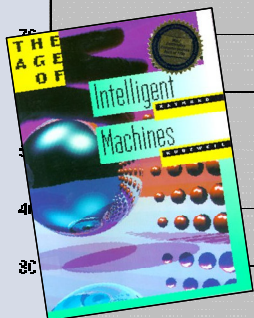
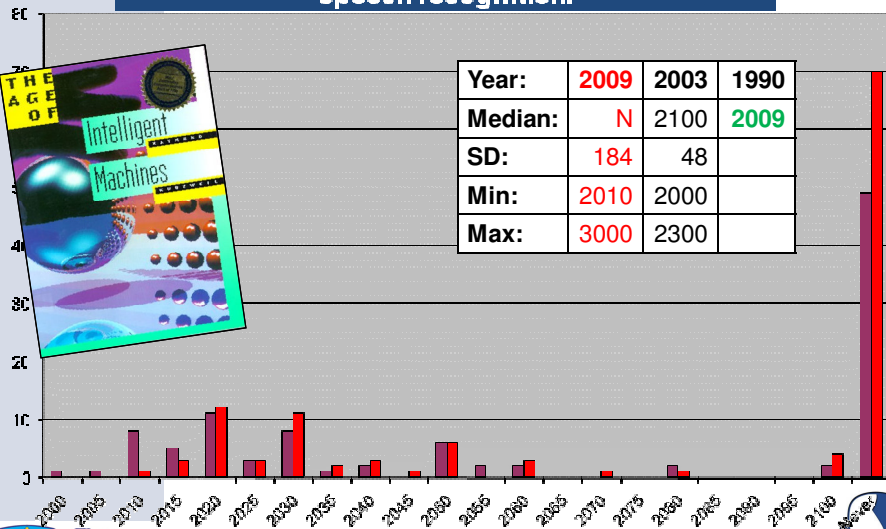
**5. Automatic airline reservation by voice over the telephone is the norm.**



The University of Sheffield.



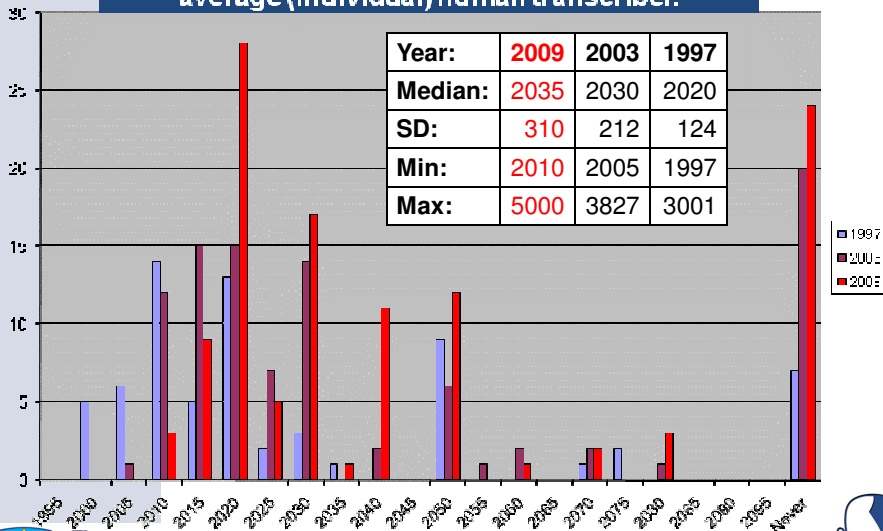
**13. The majority of text is created using continuous speech recognition.**



The University of Sheffield.



**12. Speech recognition accuracy equals that of the average (individual) human transcriber.**

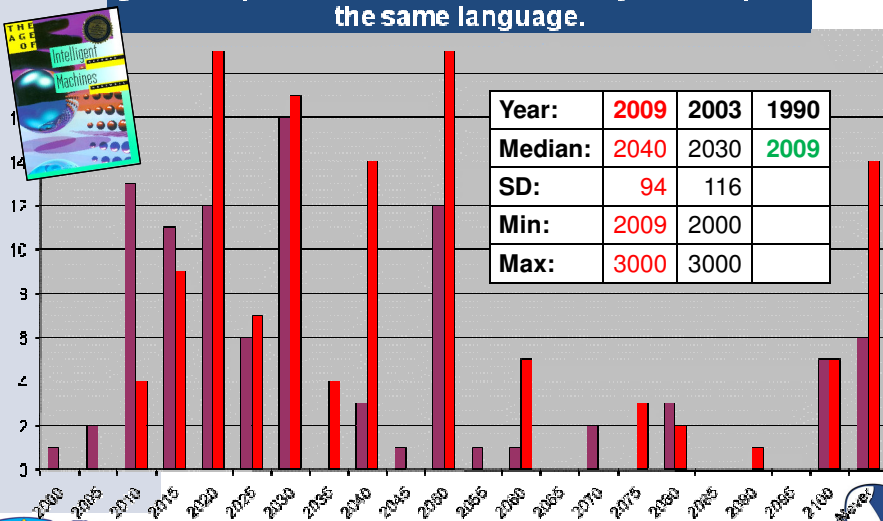


The University of Sheffield.

ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 25



**18. Translating telephones allow two people across the globe to speak to each other even if they do not speak the same language.**

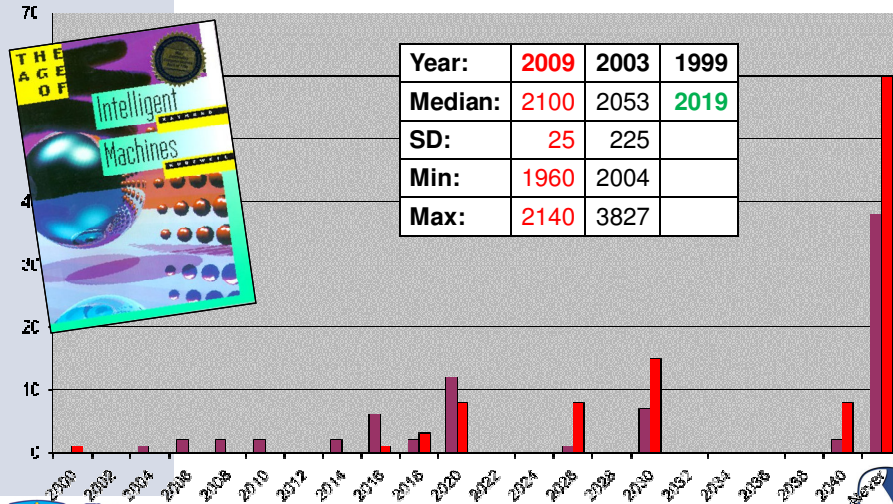


The University of Sheffield.

ASRU09 - Progress & Prospects in Speech Technology - RKM: slide 26



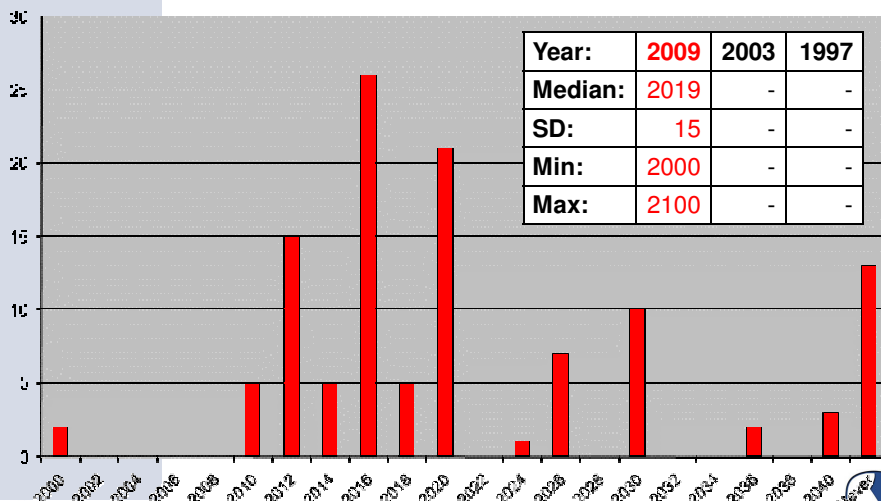
**19. Most interaction with computing is through gestures and two-way natural-language spoken communication.**



The University of Sheffield.



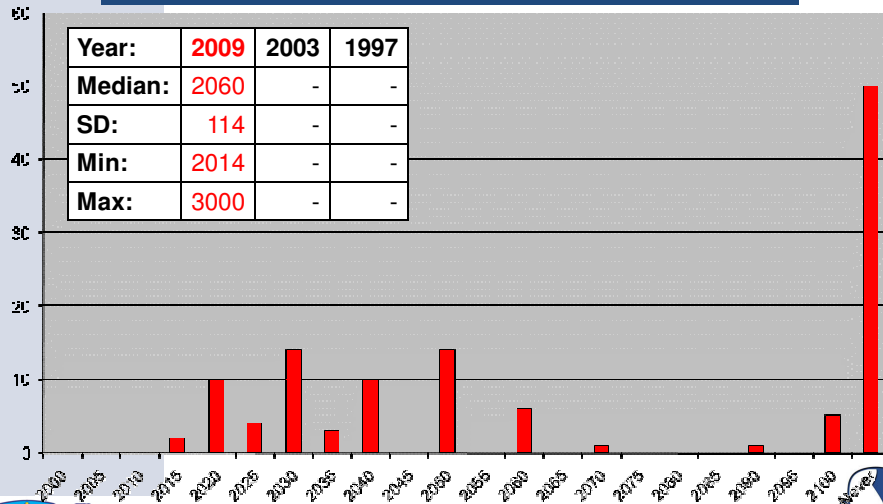
**26. All mobile devices have built-in speech recognition capability.**



The University of Sheffield.



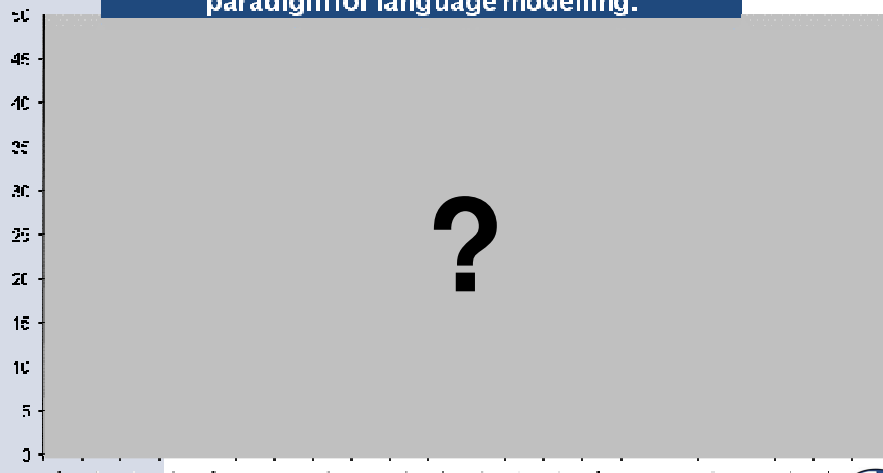
23. Most multilingual people communicate with each other through speech to speech translation at any time using their mobile device.



The University Of Sheffield.



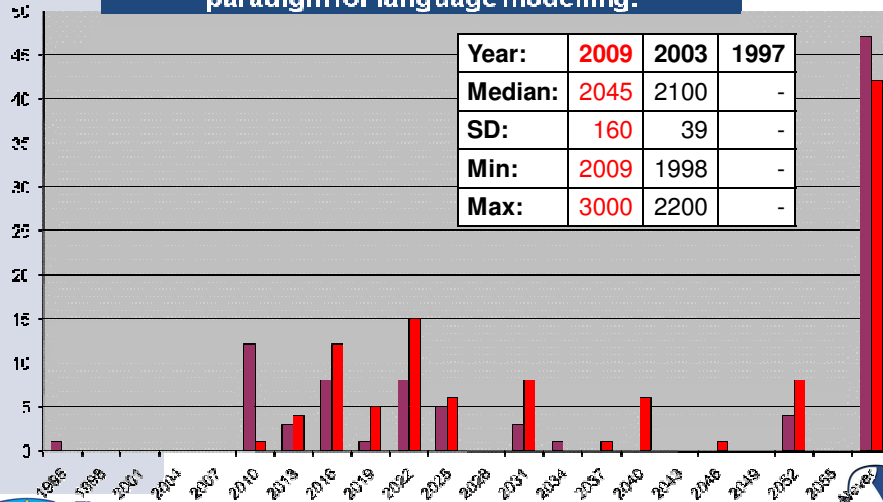
14. The majority of automatic speech recognition systems have completely abandoned the n-grams paradigm for language modelling.



The University Of Sheffield.



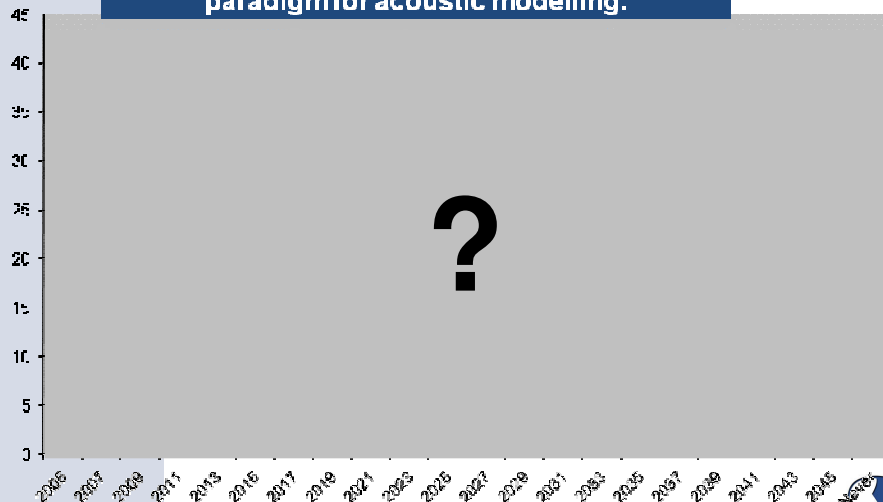
**14. The majority of automatic speech recognition systems have completely abandoned the n-grams paradigm for language modelling.**



The University Of Sheffield.



**16. The majority of automatic speech recognition systems have completely abandoned the HMM paradigm for acoustic modelling.**

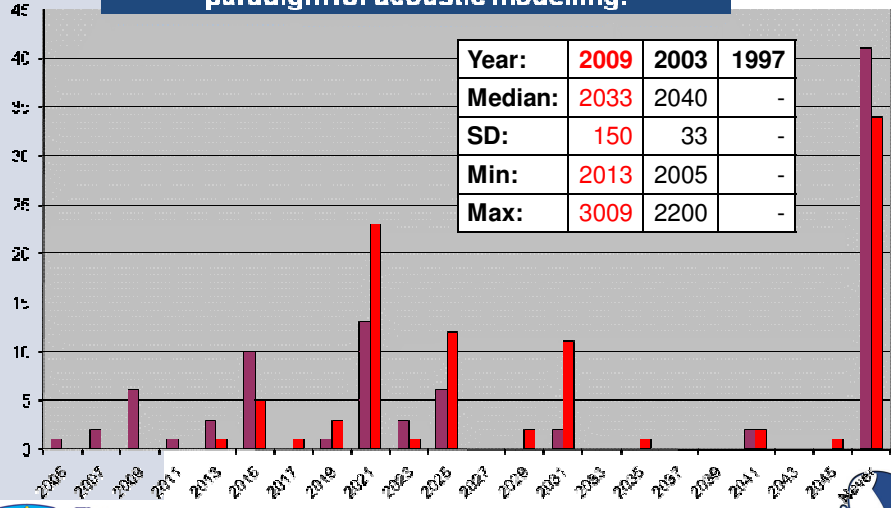


The University Of Sheffield.





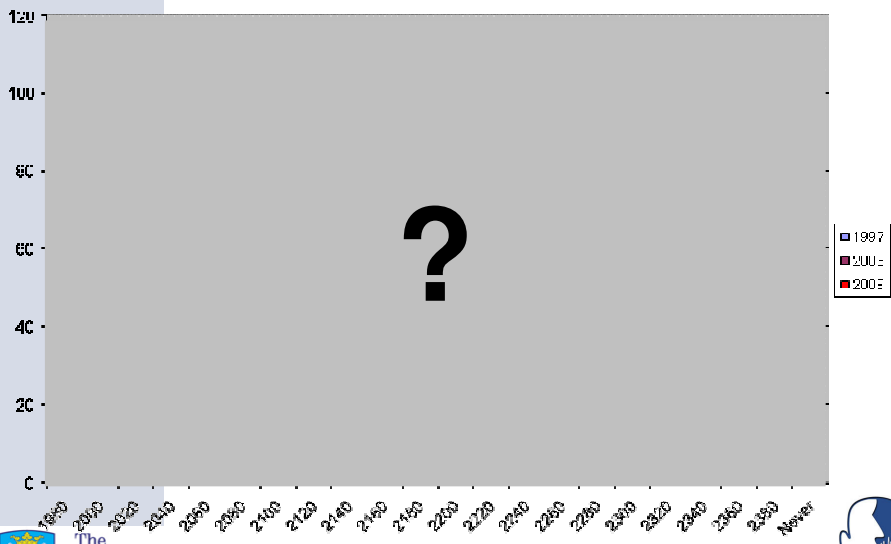
**16. The majority of automatic speech recognition systems have completely abandoned the HMM paradigm for acoustic modelling.**



The University Of Sheffield.



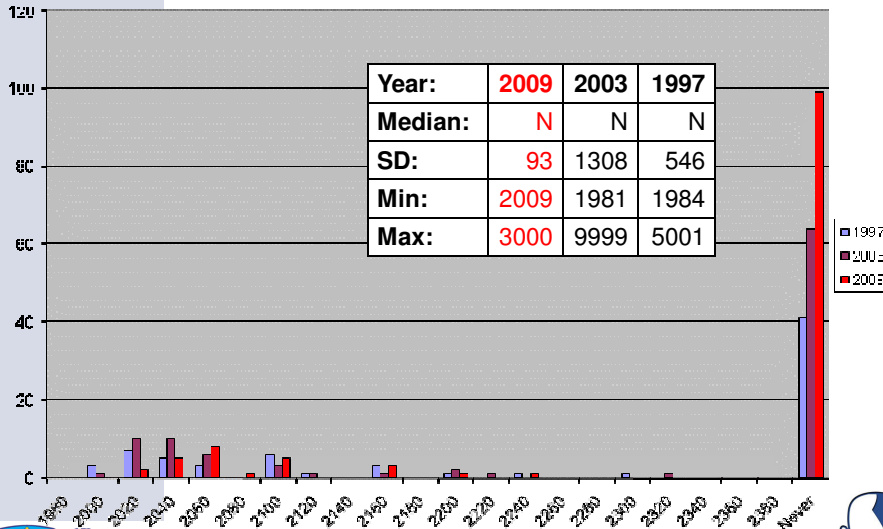
**8. No more need for speech research.**



The University Of Sheffield.



**8. No more need for speech research.**



The University Of Sheffield.



## Overall Observations

- The future appears to be no nearer than it was previously!
- The level of *scepticism* has remained remarkably stable, but *pessimism* (*realism?*) seems to have increased
- Respondents have recovered their willingness to be associated with their opinions
- The analysis reported here has been done in a rush



The University Of Sheffield.



Thank You

