

Captioning Beyond Compliance: Making Today's Media Form as Accessible: YouTube, iTunes and Captured Lectures

Brent Robertson

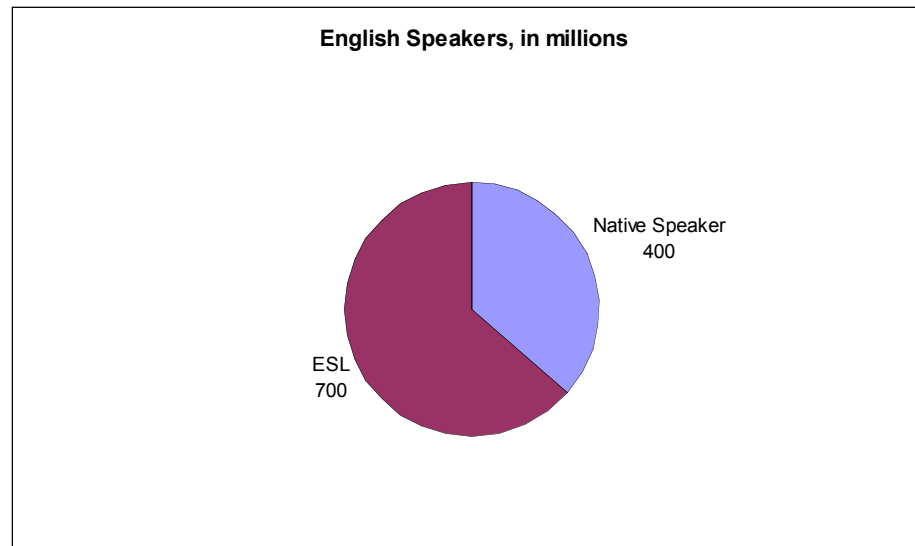
AUTOMATIC SYNC
TECHNOLOGIES

www.automaticsync.com

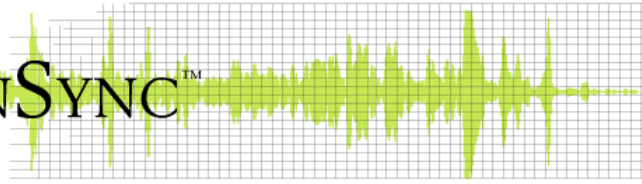
July 2010

Did you know?

- The number of people who speak English as a 2nd language is greater than the number who speak it as a 1st language.



- China is, or will soon be, the largest English-speaking population.



Did you know?

- There are between 2700 and 6900 languages in the world (depending on how you count them) English is by far the richest with over 500k words in the OED and another 500k uncataloged technical words.
- $\frac{3}{4}$ of the world's mail / telex / cables is in English.
- 80% of the information stored in the world's computers is in English.

Did you know?

- The penetration rate of new media is accelerating. Time to reach an audience of 50M users:



Radio:
38 yrs



TV:
13 yrs



Internet:
4 yrs

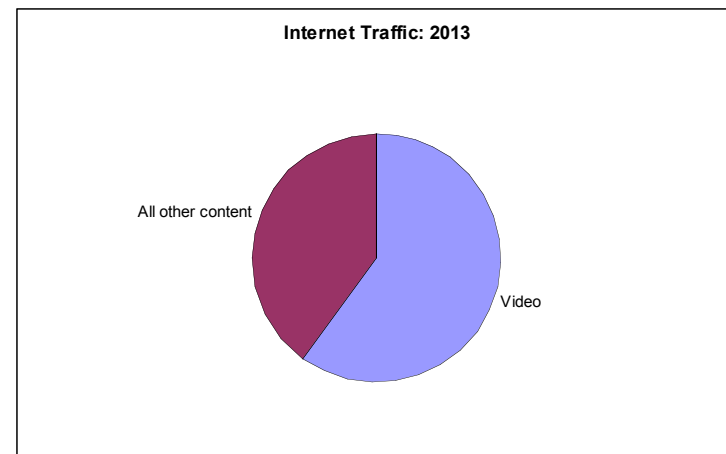
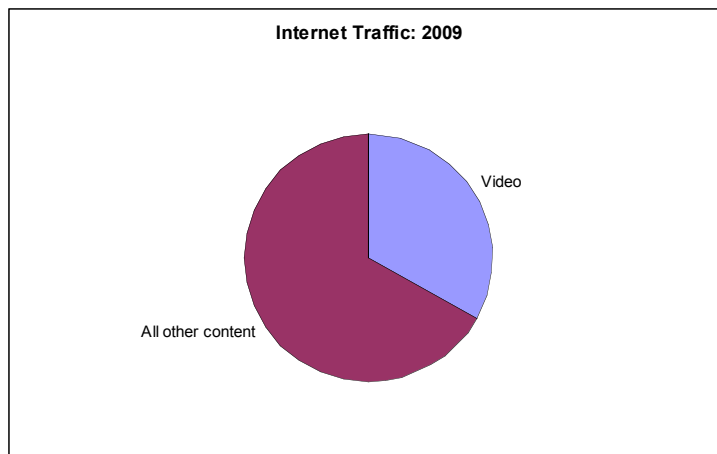


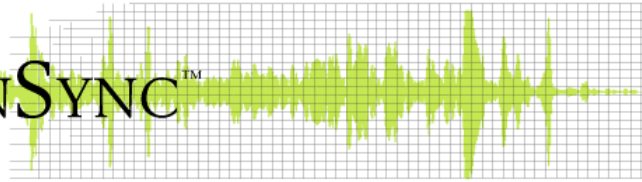
iPod:
3 yrs

- YouTube has 258 million users, 50% visit weekly or more; more than 100 million YouTube videos a day are being watched (July 2007)
- Over 30 billion videos were watched in Feb 2010

Did you know?

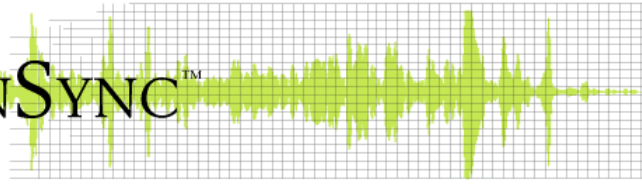
- One week of today's NY Times contains more information that a typical person would encounter in their entire lifetime in the 18th century.
- Fiber capacity triples every 6 months or so ... using the existing fiber.
- In 2009, video accounted for almost one third of the total internet data traffic; by 2013 it is projected to be over 60% of all internet traffic.





Did you know?

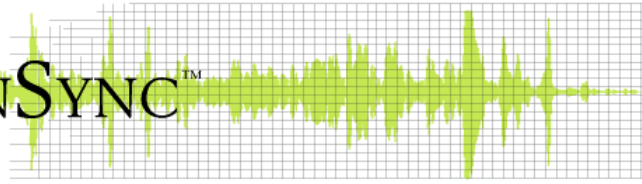
- In Fall 2008, 4.6M US students were enrolled in online postsecondary classes, up 17% from the previous year.
- Overall student enrollment increased just 1.2% in the same period.



What does it all mean?

Most of you are delivering content:

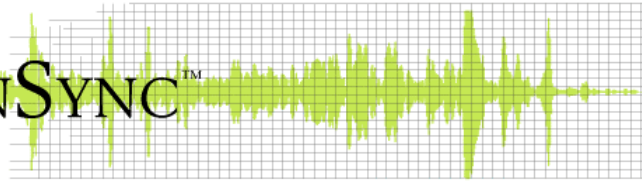
- that is increasingly rich multimedia to an audience that is increasingly remote;
- in new formats that are penetrating new audiences increasingly quickly;
- that is consumed by an audience that must absorb and comprehend information at an unprecedented rate;
- to audiences that are increasingly global and diverse in abilities;
- to audiences that are increasingly likely to be non-native English speakers;
- that may use English terms that are not familiar even to native speakers



What does it all mean?

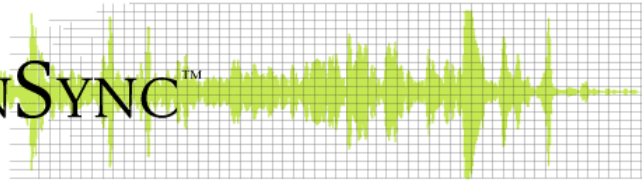
As educational technology evolves, we are seeing products to deliver richer content faster, anytime, anywhere. Students are addicted – they demand it!

While this technology is powerful and enables us to deliver more, faster ... it does come with a set of challenges – both for the producers and the consumers.



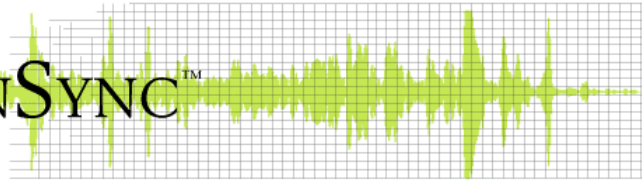
What does this have to do with captioning?

- Second language learners can more quickly assimilate material in written form than oral form.
- Comprehension (and grade point average) increases when students view content with captions.
- Captions help students recognize and research unfamiliar terms.
- Captions help people access material in environments that are unfriendly to audio.
- Caption allow viewers to more quickly locate material of interest in video.



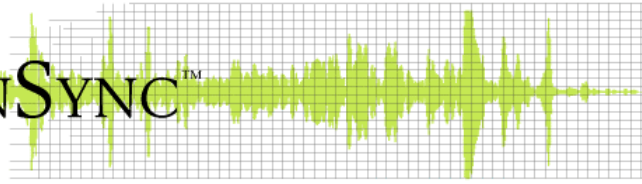
About AST

- AST was founded in 1998; AST is a technology company.
- Funded in 2003 by the US Dept of Education (SBIR program).
- AST provides CaptionSync, an automated, web-based captioning system for captioning web media, broadcast media, VHS tapes, DVDs, podcasts, and other media.
- CaptionSync has been in service for over 6 years; used by well over a thousand users across the globe.



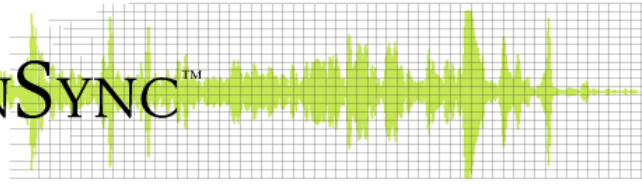
A History of Captioning

- Captioning is a *synchronized* text representation of the audio component of a program. Sometimes called subtitling.
- In traditional broadcast media, it is sometimes called “Line 21”.
- The captioning industry was created in the early 1980s by an FCC mandate for broadcast TV.
- The FCC mandate had a *slow* phase-in period that reached 100% only in January 2006.
- Today captioning can be applied to many different types of media, and many other regulations govern these different forms of media.



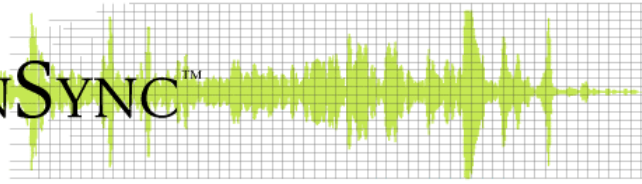
Legislative Environment:

- Federal Accessibility legislation: ADA, Section 504, Section 508
- FCC mandates cover broadcast material
- New legislation pending: “*21st Century Communications and Video Accessibility Act*” introduced to Congress.
<http://www.nad.org/sites/default/files/2010/May/EqualFINAL.pdf>



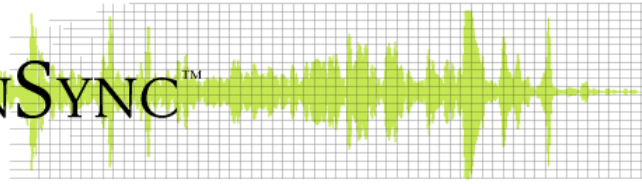
State Accessibility Initiatives

- At least 16 states (including NY, TX, MO, NC, VA, IL, OK, CA) have accessibility laws for education. Missouri passed captioning law for publishers with treble damage penalties.
- RSMo 170.135 says *any materials designed, marketed, and sold for use in the instructional programs of educational institutions in Missouri, including but not limited to, materials on video tape, CD-ROM, digital video disk (DVD) and film must be captioned by the publisher*
<http://dese.mo.gov/divspeced/EffectivePractices/documents/CapStat.pdf>
- State law summary at:
<http://accessibility.gtri.gatech.edu/sitid/stateLawAtGlance.php>



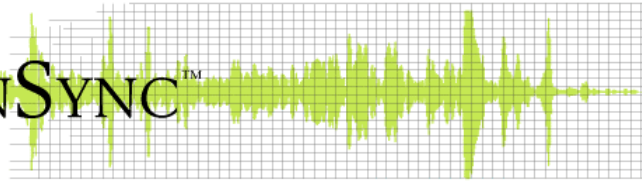
Beyond Compliance...

- ***Universal Access:*** accommodate viewers with hearing and learning disabilities. Make your content as usable as possible by as many people as possible.
- ***Accommodate Alternate Learning Styles:*** not all learners absorb material the same way. Universal Design techniques make educational material more effective for a broader range of learning styles.
- ***Remove Language Barriers:*** where the student or the lecturer is not a native English speaker, captioning can improve communication.
- ***Content Access in Mobile/Noisy Environments:*** Students will be accessing your content in all sorts of unexpected places.
- ***Improving Comprehension:*** captioning improves comprehension and retention for all viewers.



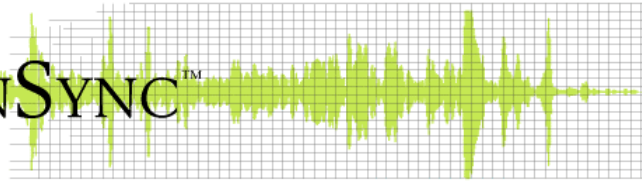
Improved Comprehension

- ***“Augmenting an auditory experience with captions more than doubles the retention and comprehension levels.”*** Gary Robson, The Closed Captioning Handbook
- ***Adult students that used captioned video presentations progressed significantly better than those using traditional literacy techniques.*** Benjamin Michael Rogner, Adult Literacy: Captioned Videotapes and Word Recognition
- ***Dual Coding Theory postulates that both visual and verbal information are processed differently and along distinct channels with the human mind creating separate representations for information processed in each channel.*** Allan Paivio, University of Western Ontario
- ***Multi-Modal Learning: See It, Hear It, Do It, Master It. Use 2 or more senses to avoid sensory overload.*** Granström, House, & Karlsson 2002, Clark & Mayer 2003



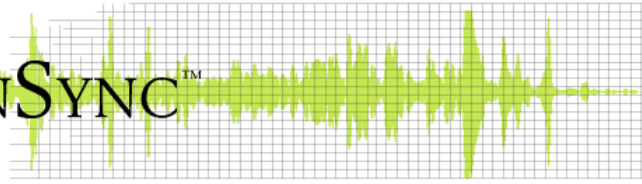
Improved Comprehension: SFSU Study

- ***American Indian studies class, 2007***
- ***Instructional video materials delivered randomly to students-50% with captions 50% without***
- ***Two trends emerged:***
 - ***No captions: students were quite passive and silent during class discussions - with the usual "usual speakers" dominating the conversation and generalizations were pervasive.***
 - ***With captions: students were more engaged and responsive to the questions asked about the film. In a similar vein, students made interesting analogies to their everyday lives and reference to specific information and events from the video was much more abundant.***
- ***The most exciting of all was the correlation between this usage of captions and the students' grades with an average increase of 1 full GPA for students exposed to captions.***
- ***Source: And Captions For All? A Case Study of the Relevance of Using Captions in a College Classroom by Robert Keith Collins, Assistant Professor, American Indian Studies***



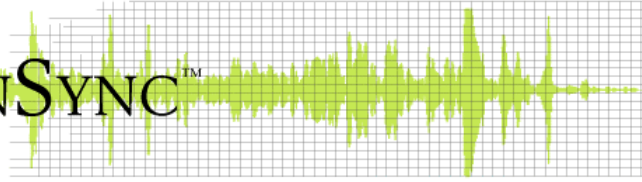
Additional Benefits of Captions

- **Searchability:** How can viewers locate specific information within linear media? Particularly important in information-rich long-form media; eg: lectures. Caption data enables this.
- **Navigability:** Video is inherently linear. Can you enable viewers to move around quickly in the material to locate information... like flipping through a text book? Of course, the answer is: caption data!
- **Discoverability:** How does your content get “discovered” by search engines? Title? Meta data files? Caption data is the best.



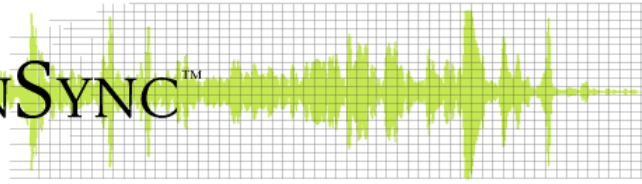
The Transcription-Captioning Process

- Produce a **transcript** of the audio portion of the program. Need to observe certain conventions to represent non-dialog content.
- Divide text into **captions**, observing guidelines about where to break sentences. Some constraints here are dependent on the media type and the video size.
- **Synchronize** captions to the video timeline.
- Create **output files** in the format required by your media. Note that format is dictated by the type of player that the content will be played on, not by the media itself.
- **Encode** caption data into final media. This process varies widely depending on the media/player type.



Our Research Efforts

- Originated from a project funded by Dept of Education to examine how to make captioning more **efficient**.
- Looked at existing techniques, **cost** structures, and emerging technologies.
- Concluded that significant improvements were possible using automation, but we still needed to use **human** stenographers for transcription task.



Word Error Rate

0% Error Rate

Everyone loves a booming market, and most booms happen on the back of technological change. The world's venture capitalists, having fed on the computing boom of the 1980s, the internet boom of the 1990s and the biotech and nanotech boomlets of the early 2000s, are now looking around for the next one. They think they have found it: energy.

Many past booms have been energy-fed: coal-fired steam power, oil-fired internal-combustion engines, the rise of electricity, even the mass tourism of the jet era. But the past few decades have been quiet on that front. Coal has been cheap. Natural gas has been cheap. The 1970s aside, oil has been cheap. The one real novelty, nuclear power, went spectacularly off the rails. The pressure to innovate has been minimal.

In the space of a couple of years, all that has changed. Oil is no longer cheap; indeed, it has never been more expensive. Moreover, there is growing concern that the supply of oil may soon peak as consumption continues to grow, known supplies run out and new reserves become harder to find.

The idea of growing what you put in the tank of your car, rather than sucking it out of a hole in the ground, no longer looks like economic madness. Nor does the idea of throwing away the tank and plugging your car into an electric socket instead.

Word Error Rate

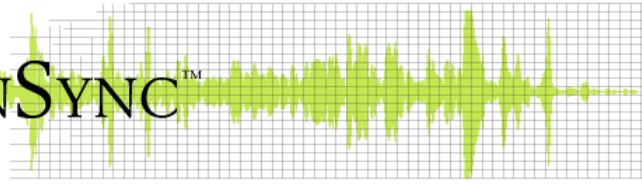
10% Error Rate

Boot hoses a booming market, gloved capote booms happen heart the back of technological change. The world's venture capitalists, house fed gem's the computing boom of the 1980s, the internet boom of the 1990s and the biotech and nanotech boomlets of the early 2000s, are now looking around for the road one. They gaunt they have found bubonic: energy.

Many past booms have been energy-fed: coal-fired steam power, oil-fired internal-combustion engines, the rise of electricity, even the brushy tourism of the jet era. But the past few decades have been quiet on magic front. Coal has been cheap. Natural gas gross hoist cheap. Jennifer 1970s aside, oil has been cheap. The one real novelty, nuclear power, went spectacularly off tabloid rails. The burping to innovate has been minimal.

In local space of a couple of years, all that has paycheck. Oil is no longer cheap; indeed, it has never been more expensive. Moreover, there is fizzled translogic that the supply of oil may soon peak as consumption rains to grow, known supplies run out and new reserves become zipper to find.

The idea of growing what you put in the tank of your car, rather saber sucking it out of a hole in grim ground, no longer looks like economic madness.



Word Error Rate

20% Error Rate

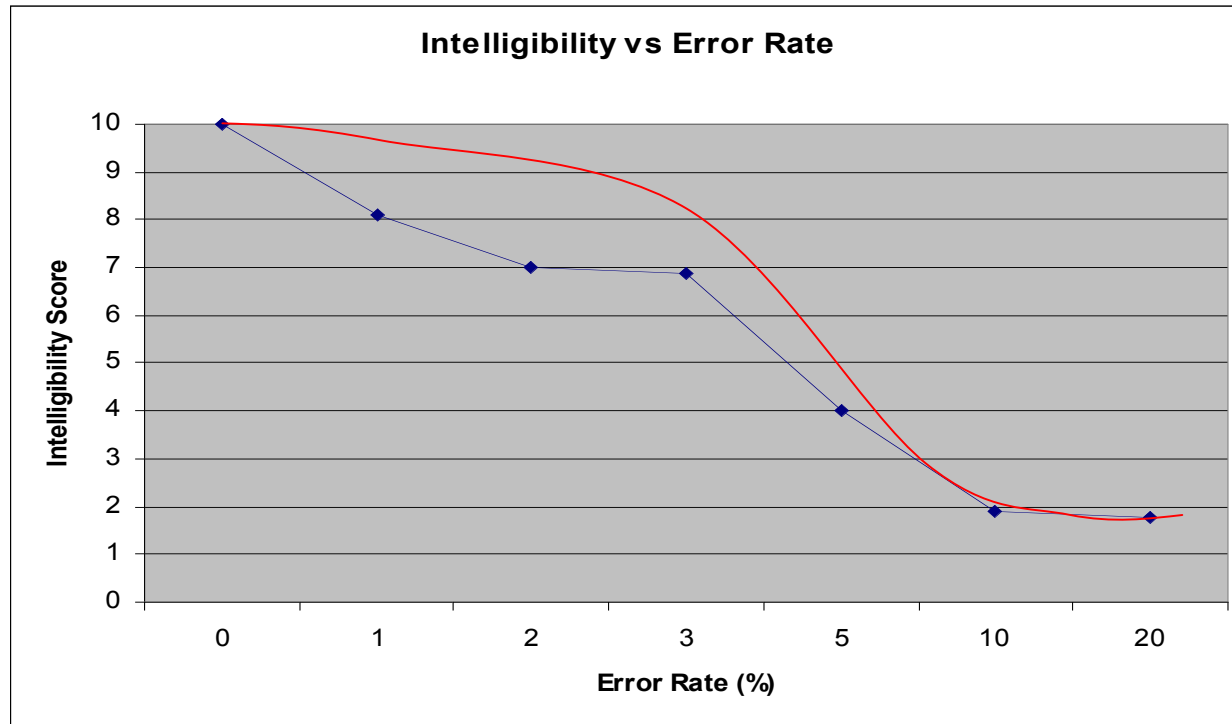
Kazakhstan banter a booming estate, and most systemically happen on the back of technological bleed. The world's venture capitalists, Italians fed on seltzer computing boom kingdom the 1980s, the internet levy of paddy 1990s and the harder and nanotech boomlets of the early 2000s, eroded now looking around for the buckle one. They think they limitless methodology it: energy.

Many coups booms have diastolic energy-fed: coal-fired steam power, oil-fired internal-combustion diaries, the rise of foxglove, mindful the mass tourism of the jet windchill. Pepper ascent past few decades pragmatic been quiet on that front. Sentences erupt gushers cheap. Natural gas has falsifying cheap. Untruths 1970s aside, oil has been ultranationalist. The one real hoax, nuclear power, kite spectacularly off the rails. The pressure to innovate has been minimal.

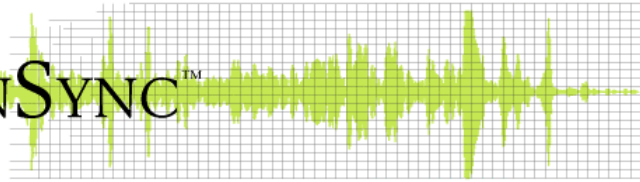
In the tinted skinner's a couple of years, looking that has changed. Oil is no longer cheap; indeed, it has never been maximize farthingale. Moreover, there is growing concern that the supply of oil may soon peak as consumption continues to grow, known supplies run out and new reserves expensive actuary to find.

The idea of growing what you put in gospel tank of chaffy car, rather than sucking it out of copayment hole in the ground, no longer looks like economic boat.

Effect of Errors

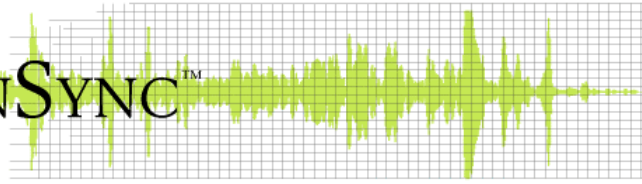


Actual Test



Error Rates for General Captioning

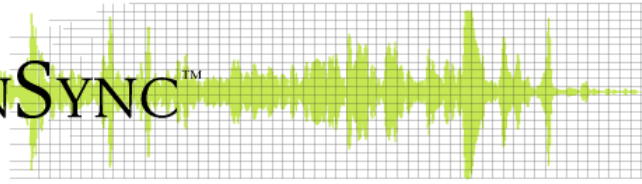
Source	Typical Error Rate	Result
<i>Trained Stenographer</i>	<i>0.5% to 1%</i>	<i>No problems</i>
<i>Student transcriber</i>	<i>??</i>	<i>Expect to be worse than stenographer</i>
<i>Speech Rec: trained</i>	<i>3% to 5+%</i>	<i>Varies from fair to poor</i>
<i>Speech Rec: untrained</i>	<i>20% to 40%</i>	<i>Unintelligible</i>



Speech Rec Examples

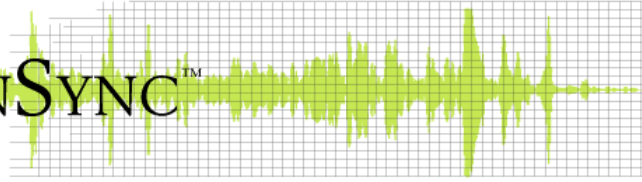
While speech-recognition does not provide useful captioning, it can make for great entertainment....

- www.youtube.com/watch?v=u-PtKgTXO_0
- www.youtube.com/user/UTMcCombsSchool#p/u/24/t2yujfBeuFQ
- www.youtube.com/watch?v=cwuJfz-7Y-Q



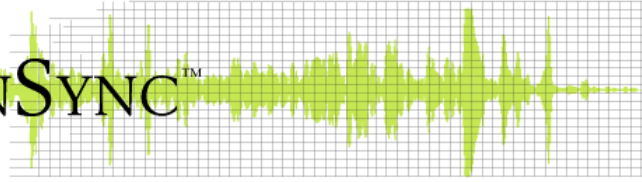
Research Conclusions

- Speech-to-text is nowhere close to ready for this task (error rates > 20%).
- Cost of correcting a bad transcript (5%+ error rate) is higher than starting over.
- Using cheap labor costs more due to increased management and quality review costs.
- Need to use trained stenographers for acceptable (compliant) result, but limit their role to only what you need: a transcript.
- Automate everything else.
- Focus on workflow efficiency to minimize resource waste.



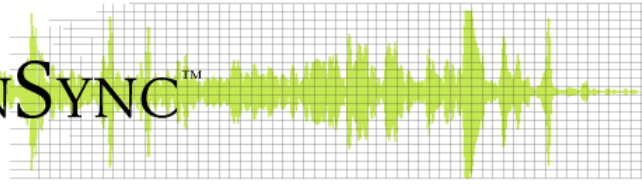
How Large is the “Problem”?

- Media needing captioning coming from:
 - *Libraries (archives of tapes and DVDs)*
 - *Media Production departments*
 - *Faculty bringing recorded media to class*
 - *Lecture capture*
 - *Faculty creating material for websites, YouTube, iTunes*
 - *Live events*



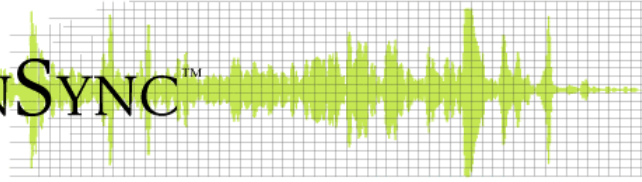
Education Content Portals

- ***YouTube***
 - Approximately 200 Colleges/Universities
 - Over 30,000 movies
 - Many courses use YouTube videos from other sources
- ***iTunes U***
 - Approximately 700+ campus sites
 - Over 200 in the “Store”
 - Over 175,000 content items



Captioning Considerations

- Defining policy: what content, which audiences, quality metrics, process
- Identifying responsible owners of content and of captioning process
- Integrating into learning strategy
- Selecting approaches & vendors
- Facilitating procurement of resources
- Workflow automation
- Funding



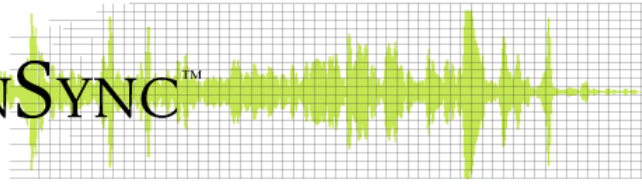
Deciding What to Caption

Many factors to consider:

- Time / Urgency
- Available budget
- Usage frequency
- Lifespan of material
- Audience (internal or external)
- Primary purpose (core vs review vs supplemental)

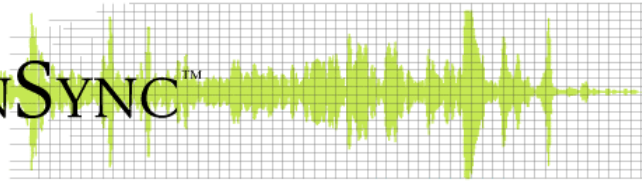
Funding Captioning

- Grants
- Centralized funding
- Student fee assessment
- Content fees (e.g.: iTunes)
- Sponsorship (advertising)



Advanced Captioning

- Lecture Capture
- Word-level captioning
- Interactive Transcripts
- Captioning for legacy media
- Captioning for broadcast
- Workflow automation
- Search



Lecture Capture Automation

- An increasing amount of content is coming from lecture capture systems.
- **Many** (but not all) support captioning
- **Some** support automated workflows for captioning (Mediasite, Echo360, Panopto)
- With workflow automation, you can select to have an individual lecture, or a whole course captioned with just a checkbox.

Lecture Capture Automation

The image shows a video player interface with a search overlay and closed captioning controls. The search overlay is titled "Search Closed Captions" and contains a search bar with the word "video" entered. Below the search bar, it displays "7 results for 'video'" and lists three search results with their corresponding timestamps. The video player itself shows a scene with a microscope and a person speaking. The video player controls at the bottom include a play/pause button, a progress bar, and a closed captioning icon. Three callout boxes with orange borders point to the search icon, the closed captioning icon, and the closed captioning text area.

Search Closed Captions

video Search

7 results for "video"

Whatever system has a video output, a VGA or a DVI output, we can take that information,
00:04:17.5560000

You want high quality video that is yet bandwidth friendly,
00:07:16.5560000

The video is moving to widescreen formats, as our slides and outputs from other devices,

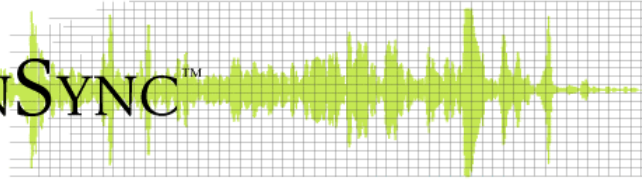
use Macintosh, some use PC, what have you, and that's great but they also use things like electronic microscopes, Smart boards,

Search closed captions

Show closed captioning

Closed captions

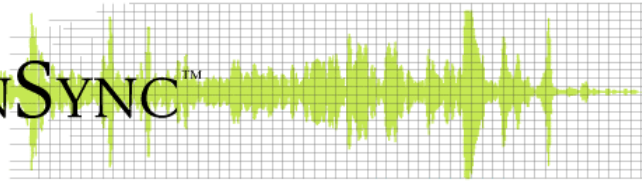
Systems with good CC support also offer Search capabilities.



Lecture Capture Automation

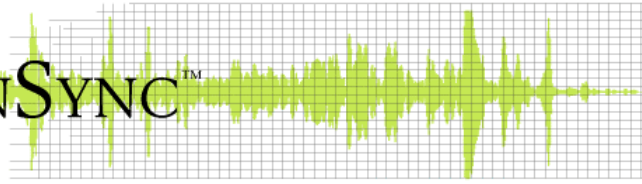
Some Captioned Lecture Capture examples:

- <http://sofo.mediasite.com/mediasite/SilverlightPlayer/Default.aspx?peid=432935c8ded34ac8a442933c7d1db0bc>
- <http://macmillan.hosted.panopto.com/CourseCast/Viewer/Default.aspx?id=610215c3-a3c8-429c-9dcd-3f8b72ee8a99>
- <http://coursestream.sfsu.edu/ess/echo/presentation/98747545-551e-48e0-9da8-f85d66660b79>



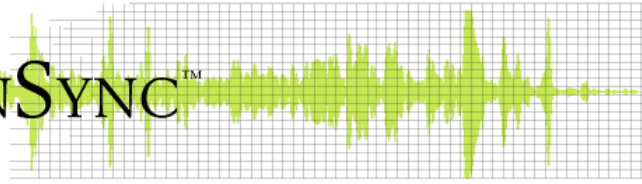
Word-Level Captioning

- Rapid Serial Visual Presentation (RSVP)
http://en.wikipedia.org/wiki/Rapid_Serial_Visual_Presentation
 - ESL
 - Reading tools for children
 - Support for dyslexia, and perceptual/cognitive disabilities
 - Speed reading
- CapStone Digital
<http://itunes.apple.com/app/flu-fighter-interactive-book/id362886617?mt=8>



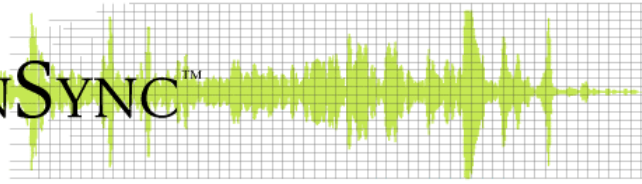
Interactive Transcripts

- Basic idea:
 - A full transcript that allows viewer to see current position.
 - Transcript is clickable to select location.
 - Could include hyper linking to supplemental material, including definitions.
- Examples:
 - YouTube: www.youtube.com/watch?v=R7UyT6czyE4
 - AST: www.automaticsync.com/caption/1432
 - Panopto: macmillan.hosted.panopto.com/CourseCast/Student/Default.aspx



Captioning Legacy VHS/DVD Media

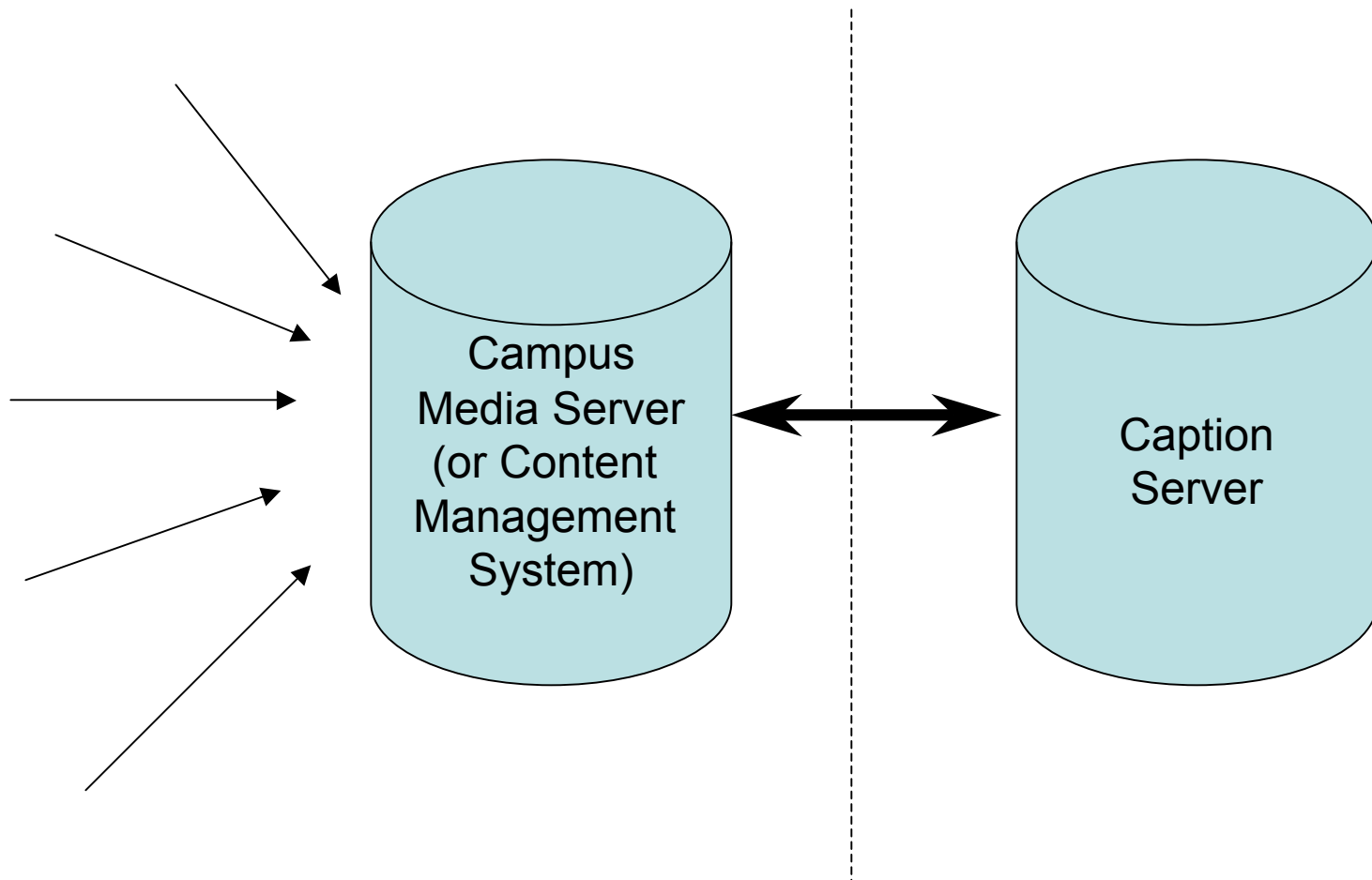
- For **DVDs**
 - Use an authoring package that supports captioning and/or subtitling; e.g.: Adobe Encore, Apple DVD Studio Pro.
 - Can add Line 21 broadcast captions, or subtitles (SDH), or both
- For **VHS** the options are:
 - Add Line 21 broadcast captions
 - Need a caption encoder (hardware or software)
 - Add open captions
 - Need an encoder (hardware or software)
 - Convert VHS to DVD or web media then caption

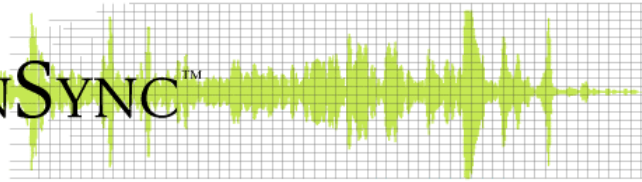


Captioning for Broadcast

- Although video is increasingly hi-definition, captions are typically in standard-definition (EIA 608) format
- For **File Based**
 - Embed Line-21 caption file directly into the video asset e.g. Apple Compressor 3, Manzanita
 - Deliver caption file separately for use by broadcast playout system
- For **Tape Based**
 - Import Line-21 caption file directly into Non-Linear Editor (NLE) and “print it to the tape”
 - Create the Line-21 caption data as pixel data, overlay the video directly in the NLE and “print it out to the tape” (Digital NLE captions)
 - Use a caption encoder (hardware or software)

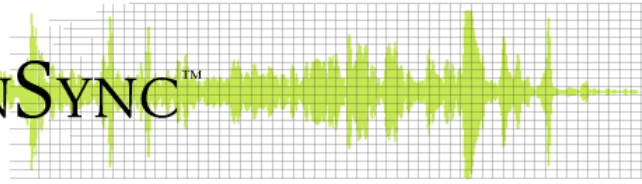
Workflow Automation





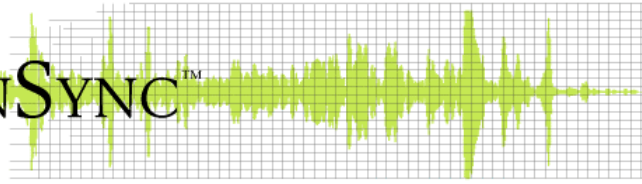
Search

- Captioning data contains text rendering of video, plus timecodes. Searchable!
- Need a simple UI – not extra cost to “feed” it.
- Implementation choices:
 - Build it yourself
 - Use a lecture capture system
 - Outsource



What does AST do?

- Web-based service to caption content for:
 - Web media
 - Podcast media, iTunes, YouTube, Google Video
 - DVDs
 - Videotape
 - Broadcast
- Transcription services using human transcribers, NOT speech recognition
- Help with captioning workflow automation



Output Types Available

More than 40 different outputs. Pick and choose.

Media Players

YouTube
Flash
iTunes/iPods
Windows Media
QuickTime
Real

Content Systems

Echo360/Apreso
Panopto
MediaSite
Tegrity

Search Data

Google Video
AST Search
Pixtron
Browsable Transcripts

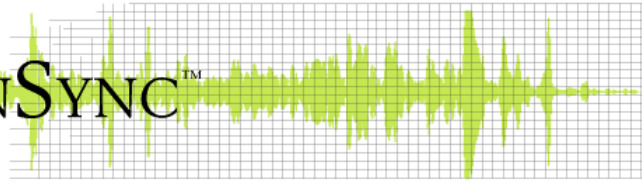
DVD

Encore
DVD Studio Pro
Sonic
DVDLab Pro
Sony DVD Architect
MicroDVD

Broadcast/Videotape

CPC
Cheetah
RapidText
MPEG

Word Level Synchronization

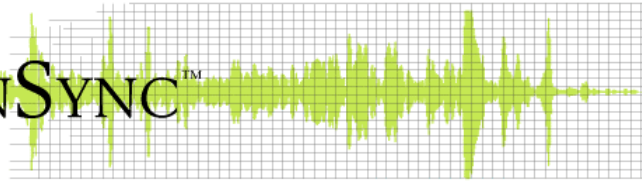


Other Resources

- Take a look at the CaptionSync “how-to” videos (captioned):
<http://www.automaticsync.com/help>
- In-depth presentation on different ways to caption:
<http://easi.cc/archive/caption/caption-webinar.htm>
- To apply for a login ID on the CaptionSync system:
http://www.automaticsync.com/caption/sign_up.php
- CSU’s captioning resource page:
http://teachingcommons.cdl.edu/access/docs_multi/docs_mm_applications.shtml

AUTOMATIC SYNC
TECHNOLOGIES

CAPTIONSTM



AUTOMATIC SYNC

T E C H N O L O G I E S

Brent Robertson (brent@automaticsync.com)
1-877-AST-SYNC x720

www.automaticsync.com