

News content

- RSS news feeds
 - Interesting News sources
 - Text news

Media Center

[Home](#) [About](#) [Arabic](#) [French](#) [German](#) [Spanish](#) [Links](#) [Search](#) [Feedback](#)

You are here: [Home](#) [About](#) [Arabic](#) [French](#) [German](#) [Headlines](#) [All Topics](#) [Politics I](#) [Politics II](#) [Business I](#) [Business II](#) [Misc I](#) [Misc II](#) [Spanish](#) [Links](#) [Search](#) [Feedback](#)

Frankfurter Allgemeine

FAZ: Politics

Parties clash over possible Turkish membership
Günter Verheugen, the German who has been the European Union enlargement since 1999, was in an optimistic mood last April. The adding 10 members on May 1, and Verheugen was letting people development the enlargement was.

A day of unity, a call for renewal
With Unification Day approaching last week, Germany's chief nun update on one key aspect of the anniversary: the number of people Germany and headed west since 1990.

Protesters turn glum
Leaders of the protest movement against the national government last week as they looked forward to a massive rally set for Saturday.

Group leader fired
The humanitarian group Cap Anamur has dismissed the organizer arrested last summer by Italian officials after he brought 37 refugees group said the decision to remove Elias Bierdeif had as much to do

die tageszeitung

TAZ
Der Gegner heißt
Globalisierung

News content

- RSS news feeds
 - Interesting News sources
 - Video news



The screenshot shows a website's 'Media Center' section. On the left is a navigation menu with the following items: Home, About, Arabic, French, Headlines, All Topics, Politics I, Politics II, **Business** (highlighted with a red box), Misc, Misc II, German, Spanish, Links, Help, Search, Feedback, Staff, Admin, Portfolio, Test, and TestVB. The main content area features a video player. The video title is 'Sharif Ali bin al-Hussein Iraqi National Congress'. The video player includes standard controls: a play/pause button, a stop button, a previous button, a next button, a progress bar with a yellow playhead, and a volume control. Below the video player, there is a text block starting with 'Dans Economia Review cette semaine : la frilosité des go...' and another block titled 'Economia Review (29.02.2003)' with the text 'Economia review revient cette semaine sur le plan de rec...'.

Content

- RSS news feeds
 - Interesting News sources
 - Audio news: *Zeit.de* with RSS



Im Zweifel für das Volk
Es ist Zeit für Plebiszite, nicht für Unterschriftenaktionen - auch beim Thema Türkei
Von Robert Leicht (6,2 MB / Abspieldauer 8 min 56)
[Real Audio](#) [WindowsMedia I](#) [mp3-Stream I](#) [mp3-Download](#)

Über Gebühr verwöhnt
ARD und ZDF bekommen mehr Geld. Ihr Programm wirft die Frage auf: Wofür eigentlich?
Von Jens Jessen (3,9 MB / Abspieldauer 5 min 33)
[Real Audio](#) [WindowsMedia I](#) [mp3-Stream I](#) [mp3-Download](#)

Wieder lieb
M... mit Besuch von Gerhard... en Kurs des libyschen...
[mp3-Download](#)

RealPlayer
Im Zweifel fuer das Volk - Von Robe... 1:50 / 8:55
8:57
Real Guide Music & My Library Music Store

EXKLUSIV FÜR ZEIT-HÖRER

€99
~~€149~~

Mobiles MP3 - Abspielgerät der Extraklasse

- Sonderpreis 99 € statt 149 €
- 128 MB interner Speicher
- Gewicht: 35 g
- bis zu 500 Minuten Abspieldauer

Content Source & Format

- RSS
 - Plain text
 - Somewhat structured

```
1 <?xml version="1.0" encoding="UTF-8" ?>
2 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3   xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:syn="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4   <channel rdf:about="http://derstandard.at/">
5     <title>derStandard.at</title>
6     <link>http://derstandard.at/</link>
7     <description>derStandard.at: Permanent aktualisierte
8     <dc:language>en</dc:language>
9     <dc:rights>Copyright (c) 2002 Ian Davis</dc:rights>
10    <items>
11      <rdf:Seq>
12        <rdf:li rdf:resource="http://myrss.com/p/f/2/t/w/gk2.html">
13          <rdf:li rdf:resource="http://myrss.com/p/c/1">
14        </rdf:Seq>
15    </items>
16    <image rdf:resource="http://myrss.com/images/logoButton.gif">
17  </channel>
18  <image rdf:about="http://myrss.com/images/logoButton.gif">
19    <title>myRSS</title>
20    <url>http://myrss.com/images/logoButton.gif</url>
21    <link>http://myrss.com</link>
22  </image>
23  <item rdf:about="http://myrss.com/p/f/2/t/w/gk2.html">
24    <title>wireless Seite</title>
25    <link>http://myrss.com/p/f/2/t/w/gk2.html</link>
26    <dc:date>2002-09-05T06:36</dc:date>
27  </item>
```

Content Source & Format

- RSS
 - Complications
 - Competing standards
 - Different encodings

```
<xsl:template match="rss:item|item|ns:item">
  <tr><td>
    <p class="rsscenterpanel"><a class="rsscenterpaneltitle" target="_blank">
      <xsl:attribute name="href"><xsl:value-of select="./rss:link|./link|./ns:link"/></xsl:attribute>
      <xsl:value-of select="./rss:title|./title|./ns:title"/>
    </a></p>
    <xsl:if test="./rss:description|./description|./ns:description">
      <p class="rsscenterpaneldescription">
        <xsl:value-of select="./rss:description|./description|./ns:description" disable-output-escaping="true"/>
      </p>
    </xsl:if>
    <hr class="rsscenterpanelitemdivider" size="0.25pt" noshade="noshade"/>
  </td></tr>
</xsl:template>
```

How to use Content: Vocabulary

- Reading comprehension
 - Which size vocabulary is needed
 - for reading newspaper articles from the target culture?
 - even to enable only context-based, global comprehension?
 - 5-6000



How to Use Content: Vocabulary

- Extensive vs. intensive reading approach
 - Definition
 - Better in promoting vocabulary growth
 - DULAP's approach: informal immersion
 - Need to adapt to students' interests
 - DULAP's approach: tailor-make careers that include foreign languages



How to use Content: Vocabulary

- Current glossing applications
 - Harvard's Text Annotator
 - Yale's CRAFT
- Not learner-centered
 - instructor glosses manually
 - texts of his/her choice
- Different Goals
 - Have computer handle large amounts of repetitive mechanical tasks (lookups, linking, tracking) to relieve humans for higher-thinking tasks
 - Allow human edits where desired



Foreign Language Data

- Dictionary data
 - Free dictionaries available (e.g. Leo.org)
 - Data need
 - to be scraped & converted into a form that can be queried
 - to be merged with word frequency data



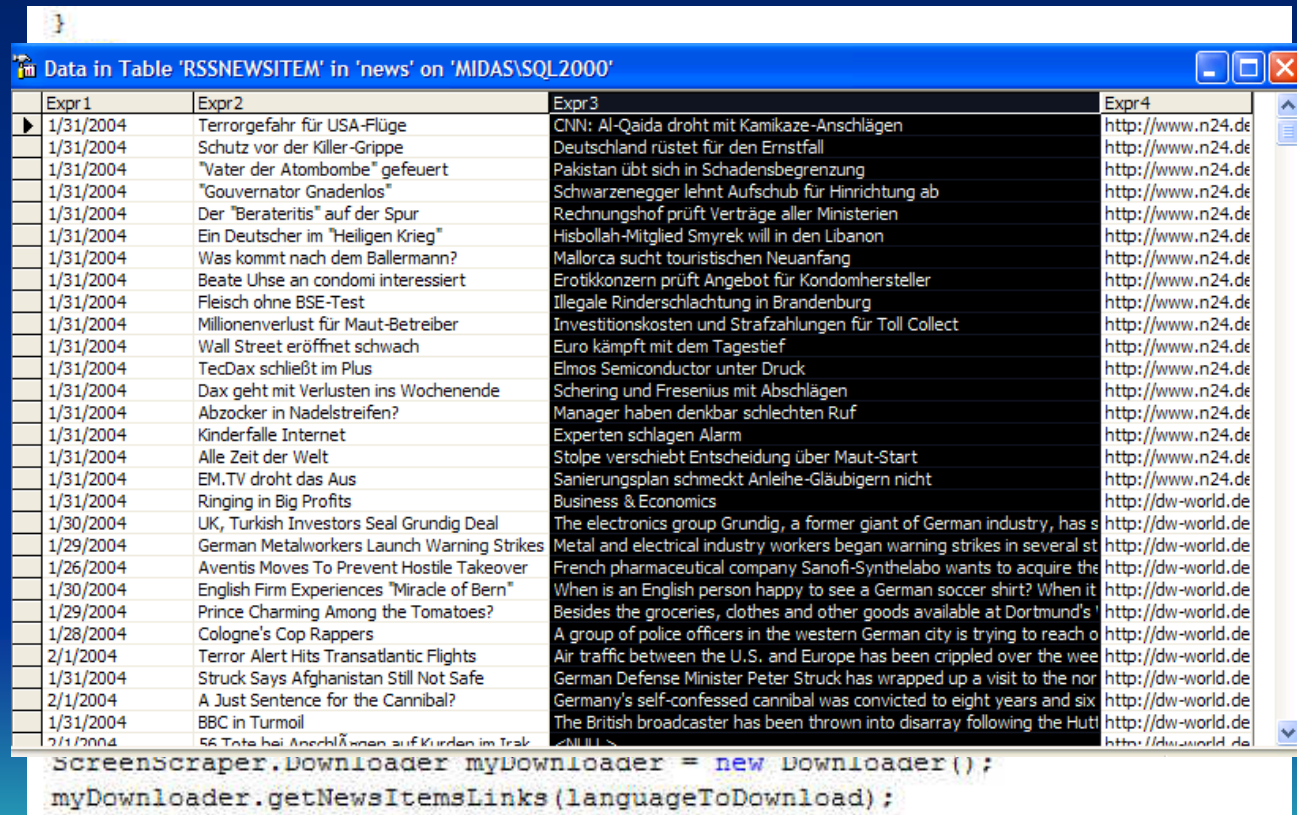
Foreign Language Data

- Word frequency data
 - Word frequency data available (e.g. [Wortschatz Leipzig](#))
 - More (and content-specific) data can be generated from teasers in RSS-item (in the database)



Import into Database

- Read out programmatically
- Store in Database for further use



Data in Table 'RSSNEWSITEM' in 'news' on 'MIDAS\SQL2000'

Expr1	Expr2	Expr3	Expr4
1/31/2004	Terrorgefahr für USA-Flüge	CNN: Al-Qaida droht mit Kamikaze-Anschlägen	http://www.n24.de
1/31/2004	Schutz vor der Killer-Grippe	Deutschland rüstet für den Ernstfall	http://www.n24.de
1/31/2004	"Vater der Atombombe" gefeuert	Pakistan übt sich in Schadensbegrenzung	http://www.n24.de
1/31/2004	"Gouvernator Gnadenlos"	Schwarzenegger lehnt Aufschub für Hinrichtung ab	http://www.n24.de
1/31/2004	Der "Berateritis" auf der Spur	Rechnungshof prüft Verträge aller Ministerien	http://www.n24.de
1/31/2004	Ein Deutscher im "Heiligen Krieg"	Hisbollah-Mitglied Smyrek will in den Libanon	http://www.n24.de
1/31/2004	Was kommt nach dem Ballermann?	Mallorca sucht touristischen Neuanfang	http://www.n24.de
1/31/2004	Beate Uhse an condomi interessiert	Erotikkonzern prüft Angebot für Kondomhersteller	http://www.n24.de
1/31/2004	Fleisch ohne BSE-Test	Illegale Rinderschlachtung in Brandenburg	http://www.n24.de
1/31/2004	Millionenverlust für Maut-Betreiber	Investitionskosten und Strafbzahlungen für Toll Collect	http://www.n24.de
1/31/2004	Wall Street eröffnet schwach	Euro kämpft mit dem Tagestief	http://www.n24.de
1/31/2004	TecDax schließt im Plus	Elmos Semiconductor unter Druck	http://www.n24.de
1/31/2004	Dax geht mit Verlusten ins Wochenende	Schering und Fresenius mit Abschlagen	http://www.n24.de
1/31/2004	Abzocker in Nadelstreifen?	Manager haben denkbar schlechten Ruf	http://www.n24.de
1/31/2004	Kinderfalle Internet	Experten schlagen Alarm	http://www.n24.de
1/31/2004	Alle Zeit der Welt	Stolpe verschiebt Entscheidung über Maut-Start	http://www.n24.de
1/31/2004	EM.TV droht das Aus	Sanierungsplan schmeckt Anleihe-Gläubigern nicht	http://www.n24.de
1/31/2004	Ringin in Big Profits	Business & Economics	http://dw-world.de
1/30/2004	UK, Turkish Investors Seal Grundig Deal	The electronics group Grundig, a former giant of German industry, has s	http://dw-world.de
1/29/2004	German Metalworkers Launch Warning Strikes	Metal and electrical industry workers began warning strikes in several st	http://dw-world.de
1/26/2004	Aventis Moves To Prevent Hostile Takeover	French pharmaceutical company Sanofi-Synthelabo wants to acquire the	http://dw-world.de
1/30/2004	English Firm Experiences "Miracle of Bern"	When is an English person happy to see a German soccer shirt? When it	http://dw-world.de
1/29/2004	Prince Charming Among the Tomatoes?	Besides the groceries, clothes and other goods available at Dortmund's	http://dw-world.de
1/28/2004	Cologne's Cop Rappers	A group of police officers in the western German city is trying to reach o	http://dw-world.de
2/1/2004	Terror Alert Hits Transatlantic Flights	Air traffic between the U.S. and Europe has been crippled over the wee	http://dw-world.de
1/31/2004	Struck Says Afghanistan Still Not Safe	German Defense Minister Peter Struck has wrapped up a visit to the nor	http://dw-world.de
2/1/2004	A Just Sentence for the Cannibal?	Germany's self-confessed cannibal was convicted to eight years and six	http://dw-world.de
1/31/2004	BBC in Turmoil	The British broadcaster has been thrown into disarray following the Hutt	http://dw-world.de
2/1/2004	56 Tote bei Anschlägen auf Kurden im Irak	CHILLI >	http://dw-world.de

```
ScreenScrapper.Downloader myDownloader = new Downloader();  
myDownloader.getNewsItemsLinks (languageToDownload);
```

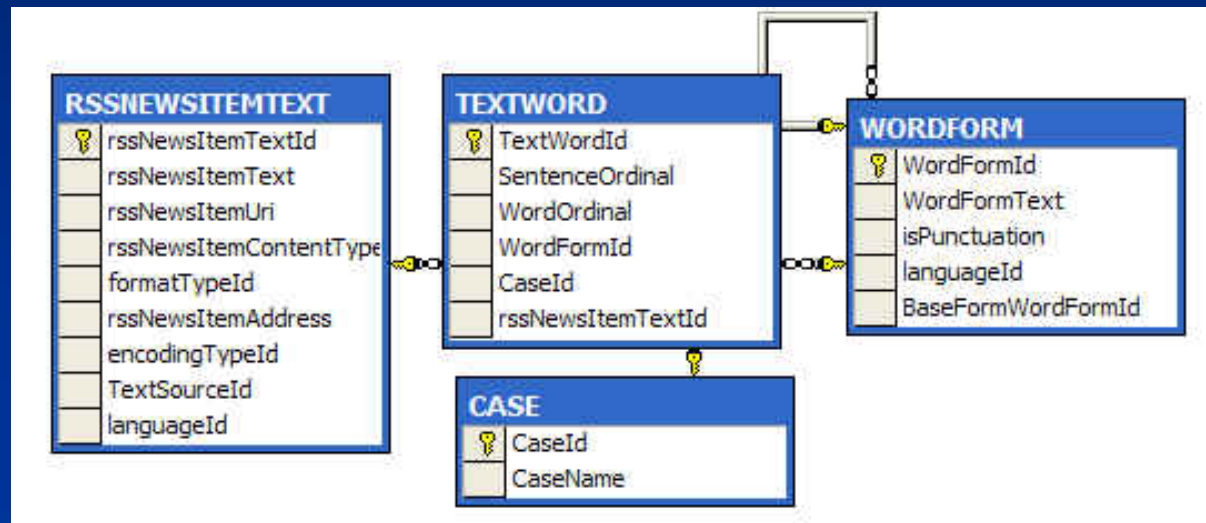
De-HTML

- Full Articles need additional processing
- Processor needs to be scriptable
- Example
 - HTML Input (.html as .txt)
 - Text Output



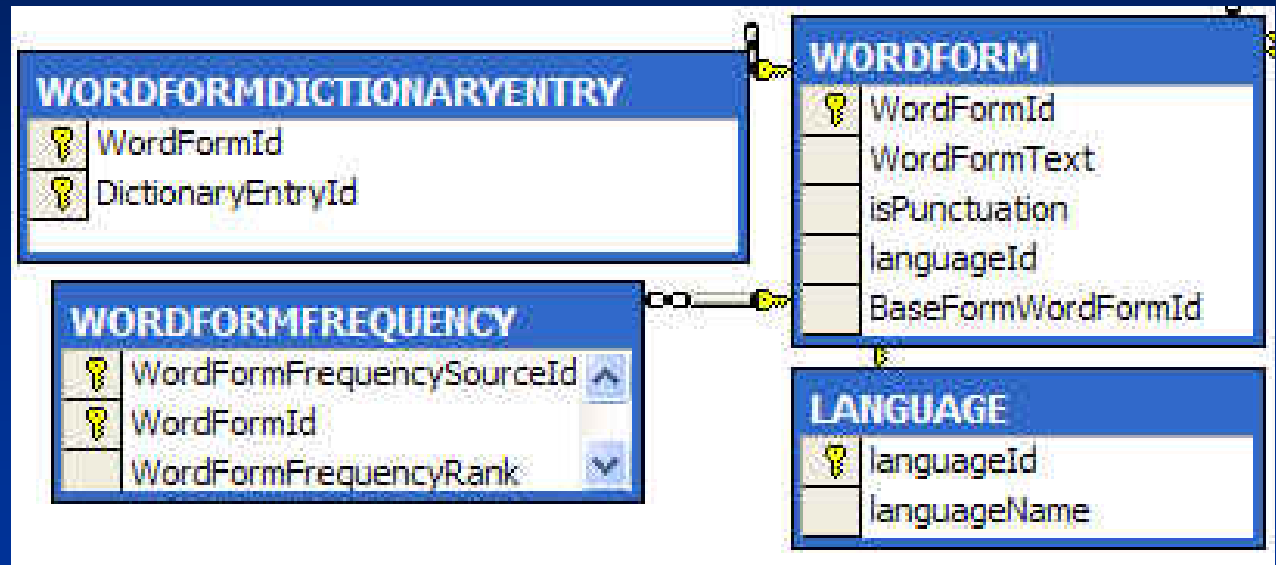
Linking

- Split headline/teaser/articles into single words, using regular expressions



Dictionary Lookup

- Lookup each word-form in dictionary-table



Computations for Foreign Language Pedagogy

- Needs a search-interface to select articles based on
 - Article length (word-number etc.)
 - Average sentence length
 - Mean word frequency
 - Generate topic-specific vocabulary lists



Deliver: Added Value

- What for
 - Focus learner attention
 - on content
 - on most frequent words
 - Frequency data allows to adapt markup to student level / interest



Deliver

- How to reassemble newspaper articles
 - As database-query into a datagrid (vertical)
 - Keep track of the wordform with wordformid
 - Add dictionary data
 - for most frequent words: with popup for less frequent words: in right column of the datagrid
 - Optionally display as synchronized captions
 - With Time+Html
 - For internet resources like Zeit.de



Deliver: Rich Text in Layers

- No “World Wide Wait”
 - For infrequent words: datagrid-Column
 - For frequent words: popups on mouseovers
 - [Blueprint Example](#)

Dies	
ist	
ein	
abgelegenes	afield adj. abgelegen , back adj. abgelegen , isolated adj. abgelegen , remote adj. abgelegen , secluded adj. abgelegen , solitary adj. abgelegen , miles from anywhere ganz abgelegen .
Beispiel	
.	

example: das Beispiel, exemplar: das Beispiel, instance: das Beispiel, sample: das Beispiel, example of: das Beispiel für, instance of: das Beispiel für, as: wie zum Beispiel

Deliver: Timed / Synchronized

- Purpose
 - Text and audio / video clip synchronization
 - Screen real estate / no scrolling
- Tool
 - HTML+Time
 - Example see next slide



Learner Interaction: Customize

- What for
 - Individualized data (learn what you do not know)
 - Include individualized context info
 - Make editable for further individualization



Learner Interaction: Customize

- How
 - Allow learner edits
 - Track learner lookups
 - And distribute to learner for further vocabulary learning
 - [Blueprint Example](#)

```
<style>
  <!-- Bind the IE timing behavior to the needed elements -->
  t\:*, tr { behavior:url(#default#time2); }
</style>
<XML:NAMESPACE PREFIX="t" />
</head>
<body>
<form>
<t:audio begin="1s" src="media/talk.mp3" syncBehavior="locked" />
<table>
  <tr begin="1s" end="4s" title="row1" onmouseover="concat = concat + 'row1'"><td>SMIL Timing
  <tr begin="4s" end="7s" title="row2" onmouseover="concat = concat + 'row2'"><td><em>attribu
  <tr begin="7s" end="9s" title="row3" onmouseover="concat = concat + 'row3'"><td>and several
  <tr begin="9s" end="13s" title="row4" onmouseover="concat = concat + 'row4'"><td>that group
</table>
<INPUT id=btnTestconcat type=button value="Which rows did I move the mouse over?" onclick="docu
```

Complications and Limitations

- Need to link word-forms to base-form
- German orthography reform
- Isolated word-approach vs. phrases
- ... [You tell me...]

