



Royal Netherlands Academy of Arts and Sciences (KNAW) KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

From a Dutch Folktale Database towards an International Folktale Database

Meder, T.

published in

Fabula

2010

document version

Publisher's PDF, also known as Version of record

[Link to publication in KNAW Research Portal](#)

citation for published version (APA)

Meder, T. (2010). From a Dutch Folktale Database towards an International Folktale Database. *Fabula*, 51(1-2), 6-22.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

pure@knaw.nl

I. Aufsätze

Theo Meder, Amsterdam

From a Dutch Folktale Database towards an International Folktale Database*

The International Society for Folk Narrative Research has several committees, and the one I am coordinating is called “Folktales and the Internet”, which deals with digitally mediated ways of interacting, storytelling, roleplaying and creating identities – ranging from SMS communication and messages on Twitter, via discussion fora and folktale websites, to being someone completely different in online ‘games’ like *Second Life* or *World of Warcraft*¹. As formulated on the website of the committee, one of the ambitions is to “create an ISFNR database with international folktale material”². In the next pages I want to make clear that such an enterprise cannot succeed without international cooperation and funding, nor without smart data-harvesting tools.

The Dutch Folktale Database

In order to get a basic idea of what such a database would look like, the existing Dutch Folktale Database can be used as an example. It has been developed at the Meertens Instituut in Amsterdam for archival reasons and as a research tool in 1994. The public version of the database that went online in 2004 can be found on www.verhalenbank.nl. There is an extended offline version of the database as well, but this can only be consulted inside the Meertens Instituut due to legislation con-

* This article gives an impression of my presentation and the following discussion at the 15th congress of the International Society for Folk Narrative Research in Athens (23 June 2009).

¹ For a further introduction to the subject see Meder, Theo: Internet. In: Haase, Donald (ed.): *The Greenwood Encyclopedia of Folktales and Fairy Tales 2*. Westport/London 2008, 489–495. Digital developments go fast, by the way: popular phenomena in 2009 like Facebook and Twitter were still to be discovered by the Dutch masses when the article was written in 2007.

² Consult <http://www.isfnr.org/index2.html> for “Folktales and the Internet” (last visit March 20, 2010).

cerning copyright (newspaper articles and press photos for instance), privacy (background information on the storytellers) and equal rights (racism, sexism).

The online database opens with a welcome page, stating that it contains over 40,000 folktales at the moment, and giving users the opportunity to make a query for stories by typing in keywords or names. It is also possible to make a more advanced search, in which there is a choice of fields to use. Any field in figure 1 (below) can be used, in any possible combination.

The database does not only contain fairy tales and traditional legends, but riddles, jokes, contemporary legends, personal narratives, and visualizations of stories as well. Most of the folktale material has been collected in the nineteenth, twentieth and twenty-first centuries, but stories from the Middle Ages and the Renaissance are present as well. Searching the database can be done by names, but the entire database is still in Dutch, so that, for example, searching for *Little Red Riding Hood* requires a search query for the name ‘Roodkapje’. The result – at this moment – is 57 hits, including *Red Riding Hood* jokes and stories that merely mention the red-hooded girl. To get more information on the first hit, clicking on ‘meer’ (more) reveals an entire page with contextual information about this version only: a Frisian variant of the *Roodkapje* story. Figure 1 gives an overview of the information presented on this page:

Unique ID
Tale type
Description of tale type
Title
Collector
Language
Source
Place of narrating
Storyteller
Date of narrating
Literature?
Subgenre
Motifs
Summary
Keywords
Names
Remarks
File name
Geographical coordinates (Kloekenummer)
Type of source

Fig. 1: Fields on the contextual page

I will give a brief explanation of the fields in use – it is not obligatory to take a look at the database, but it is helpful to understand how it works. Every story in the data-

base needs a unique ID in order to establish links with the full text, accompanying photos, sound files, movies, etc. If we chose ABIJMA22 from the list of 57 hits, we get a particular version of *Little Red Riding Hood*, told in the Frisian language. Since the database is still in the process of adapting to the ATU system, the tale type and the description are both presented as AT 0333: *The Glutton* and ATU 0333: *Little Red Riding Hood*³. Not every tale has a title – in fact, most stories taken from oral tradition do not. After all, it seldom happens that a storyteller starts a tale mentioning the title: “Little Red Riding Hood. Once upon a time, there was a little girl ...” Titles are for books. Nevertheless, the collector who made a transcription of the Frisian fairy tale added the title *Roodkapje*, although the tape recording proves that the title was not mentioned by the storyteller. In this case, the collector was Ype Poortinga (1910–85), the assistant managing director of the Frisian Academy in Leeuwarden. The language of the recorded tale needs to be characterized as ‘Woudfries’, which is a dialect of the Frisian language. The source of the text is the transcript made by collector Ype Poortinga, which belongs to the Frisian Academy. The storyteller was mole-catcher Anders Bijma (1890–1977)⁴, the place of narrating the Frisian village of Boelenslaan and the date of narrating September 28, 1971⁵. In the field that asks whether or not the source can be regarded as literature, the answer is ‘no’, because this particular recording was taped from oral tradition. The answer ‘yes’ is reserved for poems and tales taken from written sources by the collector (or read out loud by the narrator). The subgenre of Bijma’s story is ‘fairy tale’, of course – as far as this field is concerned, there is a choice from a fixed number of possibilities: exemplum, fable, stage play, kwispel⁶, song (ballad), saint’s life, limerick, joke, myth, personal narrative, riddle, traditional legend, fairy tale and urban legend. In the beginning, I used to sum up all the motifs I could find in Thompson’s Index⁷, like I did in the case of Bijma’s tale: Mot. K 2011 + Z 18.1 + F 911.3 + F 913 + Q 426. These codes indicate: *Wolf poses as*

³ For AT see Aarne, Antti/Thompson, Stith: *The Types of the Folktale. A Classification and Bibliography*. Second Revision (FFC 184). Helsinki 1961; for ATU see Uther, Hans-Jörg: *The Types of International Folktales. A Classification and Bibliography. Based on the System of Antti Aarne and Stith Thompson 1–3* (FFC 284–286). Helsinki 2004.

⁴ See Venbrux, Eric/Meder, Theo: *Anders Bijma’s Folktale Repertoire and Its Collectors*. In: *Fabula* 40 (1999) 259–277.

⁵ I must stress that a tale is linked to the storyteller in the database whenever possible. His or her performance determines where and when the story is told. This section of the database does not provide information about how old the story really is and where it originally comes from (most of the time it cannot be determined anyway). Nevertheless, I am thinking about adding a distinction between the place of narrating and the place where the narrated action is located.

⁶ A specific kind of riddles, in which the riddler gives the outcome of the story, and the riddlees have to discover the plot by asking questions. See Burger, Peter/Meder, Theo: “A rope breaks. A bell chimes. A man dies.” The “kwispel”: a neglected international narrative riddle genre. In: Catteeuw, Paul et al. [eds.]: *Toplore: Stories and Songs*. Trier 2006, 28–38.

⁷ Thompson, Stith: *Motif-Index of Folk-Literature. A Classification of Narrative Elements in Folktales, Ballads, Myths, Fables, Mediaeval Romances, Exempla, Fabliaux, Jest-Books and Local Legends*. Copenhagen 1955–58.

“grandmother” and kills child, What makes your ears so big?, Animal swallows man (not fatally), Victims rescued from swallower’s belly, and Wolf cut open and filled with stones as punishment. Filling in all these codes for every tale turned out to be extremely time-consuming and it hardly paid off in searches for certain tales: instead, tale types, keywords and names proved more helpful in quick queries. This brings us to what are perhaps the most important fields in the database: the summary offers a brief survey of the plot, while the keywords and names are very helpful in finding tales with certain motifs and themes. The keywords not only consist of the main words to be found in the full text: if necessary, certain keywords are added. For instance, if someone loses his life, the keyword ‘death’ is added. If this is because his car hits a tree, the word ‘accident’ is added. If he slits his wrists, ‘suicide’ is added, and if he gets hanged, ‘death penalty’ and ‘capital punishment’ are added. Under remarks, the person who puts the tale into the database can leave comments and explanations about the story or the performance, as well as leave remarks about additional photos, sound tracks or video recordings. The file name, ABIJMA.20E, is just the name of the backup file in Word. The ‘Kloekennummer’ is a geographical code for the place where the story was told. To facilitate his dialect research, in 1926 Gesinus G. Kloeke (1887–1963) divided the map of the Low Countries into a grid and added codes to (most) places. In fact, Bijma’s place of residence, Boelenslaan, is such a small community, that it has no ‘Kloekennummer’ of its own: B068p is the code for the nearest community Surhuisterveen – hence the question mark behind the code in the database. Finally, there is a code for the type of source we are dealing with: M stands for ‘mondeling’ which means the tale was recorded from oral transmission. Possible other types of sources can be books, CDs, stage plays, e-mails, faxes, manuscripts, the Internet, newspapers, almanacs, letters, penny prints, television, jest-books and questionnaires. Whereas a field like subgenre, for instance, enables people to just look for jokes, a field like type of source makes it possible to limit searches to newspapers only.

At the bottom of the contextual page there are a few links, which basically look like this:

All tales of this storyteller
All stories of this tale type
Geographical distribution of this tale type

Fig. 2: Links on the contextual page giving an overview

Clicking the first link shows the repertoire of Anders Bijma to the extent in which it has been entered into the database, of course. The query results in 163 hits at the moment. If a story is identified with a tale type, it is possible to click the next link and find all the stories of the same type. In the case of Bijma’s *Red Riding Hood* story, we find thirty-five tales that are of the AT(U) 333 ‘wolf eats girl’ kind. In twenty-two cases (fifty-seven minus thirty-five) Red Riding Hood is either just mentioned as a metaphor for a fairy tale, or a mere extra in a fairy-tale parody, or

the main character of some deviant joke. Thanks to the ‘Kloekenummer’, a digital map can be drawn of the distribution of the ATU 333 tale in the Netherlands. It turns out that in twenty-six cases it is clear where the story has been told.

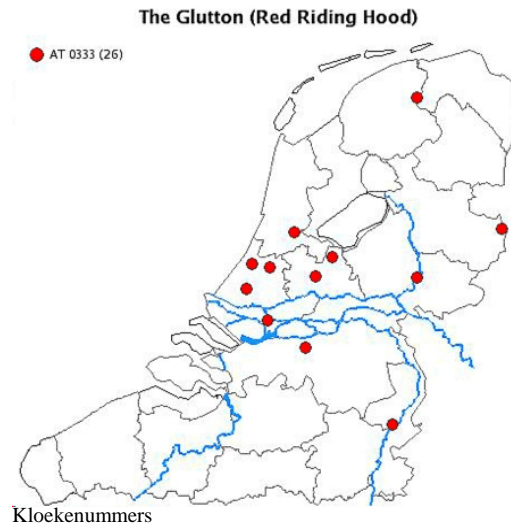


Fig. 3: Geographical distribution of 26 located versions of ATU 333 in the Netherlands

In the past, collectors interested in oral tradition were never that keen on the tale of *Red Riding Hood*, because there was a considerable chance people knew the story from written tradition. Since the collection of the tale was somewhat coincidental, nothing conclusive can be drawn from the geographic distribution in this case. Maps concerning werewolves, gnomes and nightmares are less haphazard, because in the 1960s and 1970s there was a nationwide survey of such legends.

The offline version of the folktale database has one extra link: personal information on the storyteller Anders Bijma that cannot be found on the online version. We see that Bijma was born in Boelenslaan on August 13, 1890, that he died on December 7, 1977 in Boelenslaan, and that he never went to live anywhere else. He was an orthodox Protestant and his wife was called Attje Tuinstra (1892–1980). Bijma had been a mole-catcher since 1904, and he had a small farm as well. Along the way he did some faith healing, and while trapping moles he collected archeological findings for other people. As a mole-catcher he travelled through the Netherlands, and parts of Belgium and Germany, too. Among other information, we find an interesting remark about his repertoire: Bijma himself emphasized that he got most of his stories from his grandfather (‘pake’) Alle Bijma. Just because quite a number of the storytellers in the Dutch Folktale Database are still alive today, we took all the private and personal information offline as a precaution.

Finally, there is a series of links in the right corner of the contextual page. Returning to Bijma's *Red Riding Hood* example, the next options appear:

Full Text
Grimm
Perrault
Pictures
Sound
Aarne – Thompson
Aarne – Thompson – Uther
Lexicon of folktales

Fig. 4: Links on the contextual page to other databases

The full text option is the link to the complete tale: here we find the (almost) literal transcription of Bijma's tale by Ype Poortinga. The link to Grimm offers the German *Rotkäppchen* version as a comparison⁸. Similarly, Perrault links to the French literary version, *Le petit chaperon rouge*⁹. The link pictures show an illustration from a Grimm edition, and a picture of Anders Bijma in action, catching a mole. Clicking sound leads the user to an mp3 file with the actual interview. We hear eighty-one-year-old Bijma tell the story at a slow pace, hesitating, starting off wrongly and correcting himself (he first mentions that the mother ['mem'] is ill, but corrects this into grandmother ['beppe']). It almost sounds as if Bijma had lost his ability to tell a good story, but he probably was a bit reluctant to tell this tale. In another part of the tape recording, Bijma can be heard telling Second World War stories much more enthusiastically. At the end of the *Red Riding Hood* tale, there appears to be more conversation than Poortinga's transcription suggested. Poortinga himself wrote: "Seit, dat er dit forhael op skoalle heard hat" ([he] says, that he heard this story at school). On tape, Poortinga's words can be paraphrased like this: "I don't know where you got this from, but this seems to be a story from a book, doesn't it?" We can't hear Bijma confirm nor deny this. So it looks as if Poortinga drew his own conclusions. Although he published seven volumes of Frisian folktales, he never included this *Red Riding Hood* version – or any other versions for that matter. Then there is Bijma's wife Attje: we can hear her make corrections at the end of the tale. Anders Bijma told a short and somewhat peculiar version of the Grimm tale, where the grandmother is ill, but nevertheless takes a stroll instead of being in bed at home. Red Riding Hood gets eaten by the wolf, grandma finds him and asks the hunter for help. The hunter rescues Red Riding Hood from the belly of the wolf, he puts stones back instead, and the wolf drowns when he wants to drink

⁸ For the Kinder- und Hausmärchen by Jacob and Wilhelm Grimm, I used the digital German version from the Gutenberg project: gutenberg.spiegel.de.

⁹ Transcription of: Contes de Perrault. Fac-similé de l'édition originale de 1695–1697, vol.s 1–2. Geneva 1980.

from the river. Attje says to Anders: “You didn’t tell it right”. And she starts introducing a Perrault episode, where the wolf meets Red Riding Hood in the woods first, the sick grandmother who lies in bed calls for the string to be pulled in order to open the door, and the disguised wolf repeats this formulaic phrase to Red Riding Hood. Anders Bijma doesn’t seem to be amused with this correction, saying: “Oh well then ...” Anders Bijma’s repertoire of folktales reveals him as a creative narrator, who tends to appropriate stories and make versions of his own¹⁰.

Let us now go back to figure 4: there are links to type indexes, too. The complete Aarne/Thompson catalogue from 1961 is present. In this case, the ATU 333 information appears as well, but the transition from AT to ATU is far from complete at the moment. Incidentally, the ATU catalogue, being a recent publication, is going to be a password-protected feature of the database.

Due to the fact that I took *Red Riding Hood* as an example, some catalogues in the database did not become visible. To begin with, there is the catalogue of Frisian folktales by Jurjen van der Kooi¹¹ – as Poortinga never informed Van der Kooi about the existence of Bijma’s tale and as no other Frisian versions have been taken down, AT 333 is missing in Van der Kooi’s type index. On the basis of the Frisian repertoire, Van der Kooi did create a number of new types, like for instance VDK 0958G*: *De Appelvangproef*, which in the meantime has been adapted by Hans-Jörg Uther as ATU 958F*: *Test of Sex: Catching an Apple*.

Another catalogue that is being used in the Dutch Folktale Database is the type index on Dutch fairy tales and legends compiled by J.R.W. Sinninghe (1904–88) during the Second World War¹². Although this work is old-fashioned, outdated and incomplete, it is the only catalogue we have as far as traditional legends in the Netherlands are concerned. In the database it is represented by SINUR (legends of origin), SINSAG (the *Sagen* in German) and SINLEG (the *Legenden* in German) numbers.

The contemporary legends are being catalogued in accordance with a type index that Jan Harold Brunvand made in 1993 as an appendix to his book *The Baby Train*¹³. As Brunvand was too modest to add numbers to his titles, I had to make them up myself. *The Vanishing Hitchhiker*, for instance, appears as BRUN 01000 in the database.

Still there are quite a number of tales for which the existing catalogues provide no tale type, so in order to keep track of some tales recurring in my database, I created a small catalogue of TM type numbers of my own. One example is TM 2602: *Spotnaam voor naburig dorp (stad) of hun inwoners* (sarcastic nickname for neighbouring village [town] or their inhabitants).

¹⁰ Venbrux/Meder (above, note 4).

¹¹ Kooi, Jurjen van der: *Volksverhalen in Friesland. Lectuur en mondelinge overlevering. Een typencatalogus*. Groningen 1984.

¹² Sinninghe, Jacques R.W.: *Katalog der niederländischen Märchen-, Ursprungssagen-, Sagen- und Legendenvarianten (FFC 132)*. Helsinki 1943.

¹³ Brunvand, Jan Harold: *A Type-Index of Urban Legends*. In: id.: *The Baby Train and Other Lusty Urban Legends*. New York/London 1994 (paperback edition), 325–347.

Two candidate type indexes that have not made the Dutch Folktale Database yet is Frederic C. Tubach's catalogue of medieval exempla and the catalogue of migratory legends by Reidar Thoralf Christiansen¹⁴. There was no time to integrate these works into the database yet, and I am a little in doubt whether their addition would result in many more identified folktales.

The last link mentioned in figure 4 is a lexicon of folktales. It is yet another database, but this time with background information on Dutch folktales and international tale types. The core of the database is the lexicon on Dutch fairy tales written by Dekker, Van der Kooi and Meder¹⁵. For several years now, new background information is being added to this folktale compendium, focusing especially on traditional and contemporary legends.

This completes our brief tour through the Dutch Folktale Database.

The use of having databases

The Dutch Folktale Database started in 1994 as a stand-alone database, and went online in 2004. In the year 2007 the database got visited by 1.5 million separate IP-addresses¹⁶. Perhaps this number paints a somewhat rosy picture, because in 2008 a little more than 800.000 people consulted the database. There are no exact statistics about the background of the visitors, but from the e-mails I get, I deduce that they are scientists, journalists (checking out an urban legend, for instance), students (working on an essay about fairy tales etc.), (semi-)professional storytellers (in search of repertoire), and people looking for stories their family member once told. The variety of visitors makes it clear that the database as an advanced digital archive can combine scientific, educational, artistic and social purposes. As far as I am concerned, the Folktale Database serves two major aims:

1. The archival function: The database should be an advanced digital instrument for filing and retrieving folktales. The collection of tales must be accessible on a worldwide basis and for free, and should be fit for queries by specialists and lay people. The digital database must play a part in the preservation and the accessibility of the intangible cultural heritage of the Netherlands. Many of the tales in the database are part of the unpublished archive collection of the Meertens Instituut, which has been gathering dust for several decades. At the same time, new stories are being collected during fieldwork today that need to be published in the Folktale Database.

¹⁴ See Tubach, Frederic C.: *Index Exemplorum. A Handbook of Medieval Religious Tales* (FFC 204). Helsinki 1969; Christiansen, Reidar Th.: *The Migratory Legends: a Proposed List of Types with a Systematic Catalogue of the Norwegian Variants* (FFC 175). Helsinki 1958.

¹⁵ Dekker, Ton/Kooi, Jurjen van der/Meder, Theo: *Van Aladdin tot Zwaan Kleef Aan. Lexicon van sprookjes: ontstaan, ontwikkeling, variaties*. Nijmegen 1997.

¹⁶ Meder, Theo: *Jaarverslag Documentatie en Onderzoeks-Centrum Volksverhaal juli 2007 – juli 2008*. Amsterdam 2008, 36.

2. The research function: The database should be a scientific research tool for historical and contemporary comparative folk narrative studies. In this respect, the database is not a goal in itself but a very useful device. The database is a resource for researching diachronic and synchronic variability, as well as shifting motifs and themes, moral values and beliefs, identities and repertoires of groups and individuals, and phenomena like appropriation of tales, intertextuality and intermediality. The collection of tales may in due time even serve cognitive science in researching how the human brain identifies tales and decides whether or not they are related, on the one hand, and how the computer would analyse and cluster a collection of tales, on the other. It should be very interesting to compare the results of both methods.

My scholarly colleague, anthropologist Eric Venbrux (now a full professor at the Centre for Thanatology of the Radboud University in Nijmegen) recently published an article on a narrative repertoire of Frisian tall tales concerning graveyards¹⁷. The article deals with legends and tall-tales with a Christian message that circulated in a Protestant community in the nineteenth and twentieth centuries, one of the key tales being ATU 1676 B: *Frightened to Death* (formerly known as AT 1676 B: *Clothing Caught in Graveyard*). This article was almost entirely based on material that can be found in the Dutch Folktale Database, a fact that can be regarded as a compliment for the richness and accessibility of the narrative resources and metadata offered: almost everything needed for such an article on Dutch folktales can be found in the database already.

Another thing that needs to be noted is that the data in the Dutch Folktale Database are going to be enriched with extra metadata so that it will be possible to search and harvest even larger text corpora on an international scale and participate in even more extensive scientific research. A large-scale pan-European scientific project is started called CLARIN (Common Language Resources and Technology Infrastructure), which is developed to facilitate linguistic research, but can play a role in text-based research within the Humanities and Social Sciences as well¹⁸. Tagging the text corpora including all the metadata will facilitate multi-disciplined research, and will make it more easy to combine and analyse data from multiple digital sources. CLARIN offers technology and infrastructure to access the multi-lingual sources through the Internet. I quote part of the mission statement from the CLARIN website:

“CLARIN is committed to establish an integrated and interoperable **research infrastructure** of language resources and its technology. It aims at lifting the current fragmentation, offering a stable, persistent, accessible and extendable infrastructure and therefore enabling **eHumanities**.

¹⁷ Venbrux, Eric: Spoken op het kerkhof: verhalen rond de dood uit de Friese Wouden. In: Exalto, John/van Lieburg, Fred (eds.): Spoken op het kerkhof. Verkenningen van protestantse vertelcultuur. Jaarboek voor de geschiedenis van het Nederlands Protestantisme na 1800, vol. 17 (2009) 121–134.

¹⁸ See www.clarin.eu.

- **integrated:** the resource and service centres are connected via Grid technology and form a virtually integrated domain
- **interoperable:** the resources and services will be based on Semantic Web technologies to overcome format, structure and terminological differences
- **stable:** the resources and services are offered with a high availability
- **persistent:** the resources and services are planned to be accessible for many years so that researchers can rely on them
- **accessible:** the resources and services are accessible via the web; different access methods and training possibilities are offered tailored to the needs of the communities making use of them
- **extendable:** the infrastructure is open so that new resources and services can be added easily.

[...]

CLARIN is offering its services to

- the different communities of **linguists** to optimize their models and the tools to the benefit of all using language material
- the **humanities** scholars in the broad sense to facilitate access to language resources and technology
- the **society** to enable lower thresholds to multicultural and multilingual content.”¹⁹

CLARIN appears to be turning into a promising project. Without any doubt it will enable advanced folktale research, but it remains to be seen whether folk narrative research can be one of CLARIN’s core business activities.

In general, I would like to remark that both CLARIN and my own database are still mainly fixed on (written) texts, whereas storytelling consists in performances that should be heard and seen rather than (just) read. It is true that today more and more literal transcriptions are being made of recordings from storytelling sessions, but still folktale research traditionally relies heavily on written texts alone, and the visual and auditorial aspects of storytelling are still being underestimated. In the future, databases should be able to play sound and video recordings, along with literal transcriptions according to standardized protocols and enriched with metadata in order to make the material searchable. Contemporary psycholinguistic software like ELAN and phonetics software like PRAAT make it possible to simultaneously represent several tracks as if it were a score: video, sound, transcription, annotation and tags can run over the screen in viewers synchronously²⁰.

Moreover, another genre should be taken into account when we look at narrative art forms. Sometimes texts are being replaced by pictures that are being distributed – in the past by means of the xerox copy machine, today abundantly through e-

¹⁹ <http://www.clarin.eu/external/index.php?page=about-clarin>.

²⁰ ELAN is developed by the Max Planck Institute for psycholinguistics in Nijmegen; see <http://www.lat-mpi.eu/tools/elan/elan-description>. PRAAT is developed by the department of phonetic sciences of the University of Amsterdam, see <http://en.wikipedia.org/wiki/Praat> and <http://www.fon.hum.uva.nl/praat> (last visit March 20, 2010).

mails and the Internet. I am inclined to characterize these manipulated digital pictures as ‘photoshop-lore’. Apart from the fact that they tell stories, what they have in common with oral tradition is that most of the time the maker is unknown, the pictures are distributed anonymously from one person to the other, and variations and parodies are being made on the most successful pictures²¹. Shortly after September 11, 2001, disaster jokes and riddles were told, like this one:

“Q. Bush and Bin Laden are playing chess. Who will lose?
A. Bush, because he already lost two towers.”

The same joke was visually represented and distributed like this²²:

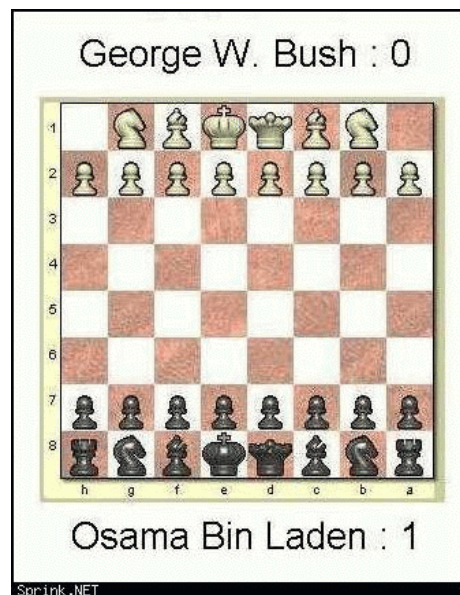


Fig. 5: Visualization of a joke, circulating in e-mails and on the Internet after 9/11

²¹ See Kuipers, Giselinde: Media Culture and Internet Disaster Jokes: bin Laden and the Attack on the World Trade Center. In: *European Journal of Cultural Studies* 5 (2002) 450–470. On p. 468, Kuipers concludes: “Interestingly, the ‘do it yourself’ style of the Internet [...] resonates nicely with traditional oral culture, which includes oral joke culture. Like the jokes and stories of oral culture, Internet jokes have no authors [...]. They are constantly created, adapted and recreated.” Furthermore, see Meder, Theo: “Welcome to Ollandistan”. *PhotoShop-lore and the Growing Perception of Division between ‘Us’ and ‘Them’ in the Netherlands*. In: Hose, Susanne (ed.): *Minderheiten und Mehrheiten in der Erzählkultur*. Bautzen 2008, 259–277.

²² Both jokes can be found in the Dutch Folktale Database under the ID-numbers DVTEX124 and HUMOR296.

I believe a modern folktale database should not only contain folktales from written and oral sources, but visual material like this piece of Photoshop-lore as well, while some Powerpoint presentations and movies can appear as folktales as well. A folktale database should nowadays be a multimedia database, containing not only text, but sound, pictures and movies, too.

An International Folktale Database

The initial idea of an International Folktale Database was launched by Rolf Wilhelm Brednich in an ISFNR workshop on computer-mediated communication in Tartu (Estonia), in the summer of 2005. Not only did he plead for the narrative study of digital texts and pictures, he also was in favour of the development of an international database, especially for the digital narrative material. As we all have experienced, the Internet can be as volatile as oral tradition: a story can be here today and gone tomorrow. The same goes for the websites with folktales, mentioned in the footnotes of articles in *Fabula*: websites are not as reliable as books, because they tend to change, move or even vanish into cyberspace. Footnotes in a scientific journal could therefore turn worthless within a few years. The solution would be an International Folktale Database in which webpages with folktale material remain preserved for a long time.

Since I became the coordinator of the ISFNR committee for Folktales and the Internet, I volunteered to have a prototype of this international database built and host it on a server in Amsterdam. In the summer of 2009, during the ISFNR congress in Athens, I gave a demonstration of both the Dutch Folktale Database and a password-protected demo-version of the International Folktale Database, which I filled with six stories from several languages: English, French, German and Dutch²³. Apart from the texts, which are still in their original languages, the fields of the IFD are in English now. Most fields have remained the same. Apart from the place (of telling), the name of the country needs to be filled in now, as well as the coordinates, which will enable a geographical representation in Google Maps (or even Google Earth). Furthermore, a field called reference is added, to be able to refer to a note in an issue of *Fabula* in which a folktale website is mentioned.

As a first example, DEMO001 was shown: a cartoon taken from Brednich's book *www.worldwidewitz.com*²⁴. Under the image button, the cartoon can be found. The text itself is still represented in German: "Was haben Sie denn gedacht, wie er aussieht?" The summary of the plot is in English, though: "A cartoon of Saint Peter, showing God to an angel: God is a computer. Saint Peter says: "What did you think

²³ See <http://www.meertens.knaw.nl/ifd>; password can be provided on request (mail to Theo.Meder@Meertens.knaw.nl).

²⁴ Brednich, Rolf W.: *www.worldwidewitz.com. Humor im Cyberspace*. Freiburg/Basel/Wien 2005, 123.

he looked like?” The keywords are in English: “angel, god, computer, heaven, show”, as are the names: “Petrus, Saint Peter, God”.

The next example (DEMO006) was a bilingual Powerpoint joke (a slideshow with text and pictures) from 2006, featuring soccer player Ronaldo, US senator and former president’s wife Hillary Clinton, president George Bush, pope John Paul II and a little schoolboy. Their plane is crashing, and they are one parachute short. The pope wants the kid to have the parachute, but there is no need: Bush just jumped down with the boy’s school pack on his back. The last pictures of the powerpoint presentation, which can be downloaded under the powerpoint button, make it clear how much the joke intends to portray Bush as dumb. The powerpoint joke, forwarded by researcher Véronique Campion-Vincent from Paris, is a typical example of a tale distributed by e-mail.

After the cartoon and the powerpoint joke, there came a commercial film advertising rain coats (DEMO005): a couple hits a kangaroo with their car in the Australian desert. Because it is raining, they put a rain coat on the shaken kangaroo and then make a picture. After that, the kangaroo runs off with the man’s rain coat, the car key in one of the pockets. Under the movie button, the commercial can be found. The story is a version of a well-known contemporary legend, which researcher Jan Harold Brunvand calls *The Kangaroo Thief (The Gucci Kangaroo)*²⁵. I added a number for this specific tale type: BRUN 01120A.

The last example I gave (DEMO004) showed how to deal with a frozen webpage. In 2007 I wrote an article about crop circle belief and modern exempla in *Fabula*²⁶. In note 37 I refer to a specific webpage about a crop circle hoax on the site of the Dutch Crop Circle Archive²⁷. In the demo version of the International Folktale Database I saved this webpage as a jpeg-image screenshot: even if the webpage has been altered in the meantime, it remains preserved in the original state in which I consulted it in the past. I have added a transcription of the page (in Dutch) and all the necessary metadata (in English) as well. Now the real webpage can change, move or vanish altogether, but future readers of *Fabula* can still check the webservice mentioned in my note 37. The only problem with jpeg images is that all dynamic parts – from links to animations – in the original webpage will be frozen, too. In due time, a more sophisticated technical solution needs to tackle this problem.

I concluded my presentation in Athens with some questions, like: do we really need an International Folktale Database? What data should this database contain, and what problems will we most likely encounter? The final question was: will it work?

As far as the necessity of an International Folktale Database is concerned, the research community could easily opt for the old-fashioned methods. Most researchers

²⁵ Brunvand (above, note 13) 326.

²⁶ Meder, Theo: Modern Exempla: Crop Circle Tales in the New Age Era. In: *Fabula* 48 (2007) 281–299.

²⁷ www.dcca.nl/1998/nl42.htm to be more precise (last visit March 20, 2010).

are accustomed to work with books and manuscripts – why should we change that? I do not think that this is necessary, but going digital can provide us with much added value. Stories in archives that are difficult to obtain or even completely forgotten could be accessed worldwide for free if we decide to put them in an online database. In a world that is going digital fast, we need to start thinking about storing and making accessible our folktale material in digital databases, with the multimedia possibility to contain text, picture, sound and movie. Not only in order to safeguard our intangible cultural heritage, but also because search engines can manage much larger amounts of international folktale data for the sake of scientific research. Furthermore, I do not suppose we will carry on making paper catalogues: the next generation of folktale catalogues will be digital. The way in which catalogues and stories can be interlinked again will provide added value. Apart from this, Brednich's idea of freezing and safeguarding webpages mentioned in scholarly articles makes perfect sense to me.

The question what features are necessary to make an international database into a success is connected with the issue of how the database should be structured and how it should operate. This matter will be dealt with later on. One of the problems which will most likely be encountered is the lack of an international catalogue of traditional legends that would help researchers to discover and identify legends of the same branch circulating internationally. One of the last attempts to reach a consensus concerning the compilation of a type index of legends failed in the 1960s²⁸. On the other hand, tales can be put into the database without a type number from any of the folktale catalogues. In the case of the Dutch Folktale Database, about 40% of the folktales remain without a type number, because there is no catalogue providing any. Not only the personal narratives and, for instance, the 'kwispels'²⁹ for which there is no room in indexes are to blame for this percentage. Especially traditional legends and modern jokes are not easy to catalogue, even after I have added a small TM-catalogue for recurring legends and jokes in my own database (TM = Theo Meder). Of course, this is a problem that cannot be solved completely, because there will always be uncatalogued folktales – still, developing an internationally recognized (preferably digital) catalogue for traditional legends would be helpful. Incidentally, although I hardly use the field for motifs, I would like to plead in favour of its preservation. Most likely, in an international database it may become interesting to search for Thompson's motif codes, particularly if we could manage to interlink the entire *Motif-Index* with the database³⁰.

The final question "will it work?" presupposes at least two conditions:

1. International cooperation. Several parties like the Meertens Instituut, the Flemish Folktale Database, the *Enzyklopädie des Märchens* etc. need to work together in

²⁸ Greverus, I.-M.: Bericht zu Veröffentlichungs- und Katalogisierungsplänen aus dem Zentralarchiv der Volkserzählung. In: *Acta Ethnographica* 13 (1964) 111–128, specifically the 'Diskussion und Arbeitsergebnisse', 129–131.

²⁹ Burger/Meder 2006 (above, note 6).

³⁰ Thompson (above, note 7).

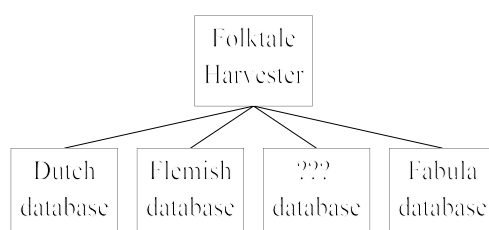
order to achieve international esteem. CLARIN may be another valued initiative. Still, cooperation from other continents is necessary as well.

2. International funding. Without funding we will not get far. Meanwhile, CLARIN may be willing to invest as long as it serves CLARIN's (European) purposes.

The need for a data-harvesting tool

During the discussion one of the most fundamental – and basically psychological – questions was raised: will a large database, hosted by one central party, work? That is: are scholars and students all over the world prepared to put their data in someone else's database? Probably not. Besides, will the host be able to support, maintain, evaluate and (if necessary) correct a vast amount of multilingual data? This would be very unlikely. The most conceivable outcome would be that every group wants to keep, maintain and be responsible for their own data – be it on the level of university research units or on a regional or national scale.

Fig. 6: The data-harvester can search any number of designated online databases



It seems much wiser to build an international folktale database harvester: a powerful search engine in a grid of designated online databases, which will be able to obtain specific data from a large number of separate digital collections. The data-harvester can be programmed to perform all kinds of searches, like a spider in a folktale web – think of it as a sort of ‘Google Folktales’. Beforehand, of course, the data-harvester needs to be instructed in which online collections to search, and in which fields specific metadata can be found. The more a certain database is uniformized to meet the harvester's requirements, the more useful data it can provide. The more fields as mentioned below in figure 7 are filled in in separate databases, the more data the harvester will retrieve after a query.

Full text (in original language)	Storyteller
Unique ID	Date of narrating
Type (ATU)	Subgenre
Description	Literary?
Title	Motifs (optional)
Collector	Plot (summary)
Language	Keywords
Source	Names
Source type	Remarks
Reference	Picture
Place of narrating	Sound
Country	Powerpoint
Geographical coordinates	Movie

Fig. 7: Required fields in the databases

The stories themselves should be provided in the original language, but the other fields mentioned in figure 7 should preferably be in English. If possible, the folktale should at least be classified as an ATU type. Keywords, names and summaries in English would be very helpful as well (although perhaps in the future translating tools could be of use to us). If fields are left open, or another language than English appears, the providing database is responsible, not the harvesting tool. Plain folktale texts which are not enriched with metadata, like the fairy-tale editions published by the Gutenberg project³¹, will not supply any extra information (like ATU types). In this case, simple text queries will only have Google-like results.

The Dutch Folktale Database will become part of the CLARIN initiative. A possible first pilot project could be to create a harvesting tool for research within the CLARIN grid, which will be able to pair and search the Dutch Folktale Database in Amsterdam and the Flemish Folktale Database in Leuven.

If necessary, the Meertens Instituut could host a database for *Fabula* or other colleagues who do not feel like creating an online database of their own. For scholars in need, the Meertens Instituut can provide folktale database resources as well.

The Meertens Instituut is prepared to play a leading role in the process of digitalization, building databases and a harvester.

Abstract

Some folk-narrative scholars have been creating (national) folktale databases, and there are probably more to come. Among the existing ones is the Dutch Folktale Database with over 40,000 folktales and metadata. If the ISFNR wants to build an International Folktale Data-

³¹ See gutenberg.spiegel.de.

base, it would be best to create a number of national databases, for instance, link them in a grid and have a smart harvesting tool perform queries. The more the separate databases get uniformized, the better the results of the data-harvester will be. This enterprise needs international scholarly cooperation and funding. The Meertens Instituut in Amsterdam is prepared to take the lead.

Résumé

Quelques spécialistes en matière de la littérature orale sont en train de développer des bases de données des contes populaires (nationales), et très probablement il y en aura davantage à l'avenir. Parmi les bases de données déjà existantes se trouve celle des contes populaires néerlandais, qui contient plus de 40.000 contes populaires et données meta. Si l'ISFNR aspire à constituer une base de données internationale des contes populaires, le plus efficace serait de créer un nombre de bases de données nationales, par exemple. Ensuite, on pourrait les connecter dans un réseau informatique et exécuter des recherches avec un outil d'exploration intelligent. Plus ces bases de données seraient uniformisées, plus les résultats de l'explorateur de données seraient précis. Une entreprise pareille aurait besoin de coopération et financement sur un plan international. Le Meertens Instituut à Amsterdam est prêt à diriger cette opération.

Zusammenfassung

Einige Erzählforscher sind bereits mit dem Aufbau (nationaler) Volkserzählungsdatenbanken beschäftigt, und ihre Zahl wird vermutlich immer größer werden. Zu den schon existierenden Datenbanken dieser Art gehört die niederländische Volkserzählungsdatenbank, die mehr als 40.000 Volkserzählungen und Metadaten enthält. Falls die ISFNR eine internationale Volkserzählungsdatenbank aufbauen möchte, wäre es am besten, zum Beispiel vielfältige nationale Datenbanken zu schaffen, die man in einem Raster miteinander verbindet. Danach könnte man Suchaktionen mit Hilfe eines intelligenten Suchwerkzeugs ausführen lassen. Je stärker die einzelnen Datenbanken vereinheitlicht würden, desto genauer wären die Ergebnisse des Datensammlungsinstruments. Ein solches Unternehmen würde internationale wissenschaftliche Zusammenarbeit und Finanzierung erfordern. Das Meertens Instituut in Amsterdam wäre bereit, eine führende Rolle dabei zu spielen.