Ukrainian Folklore Audio

Natalie Kononenko

The Internet provides scholars with an ever-expanding variety of ways to interact with the public. Crowdsourcing, or putting one's research tasks online and asking for help from volunteers, is perhaps one of the most rewarding things an academic can do. The assistance one gets is not only free labor, but it is also reassurance that others are interested in the things that we find fascinating. Crowdsourcing, as I learned from the Ukraine Folklore Audio project, also provides researchers with valuable information about the public. When I first developed the idea for this project, my goal was to have prospective users help me select which of the many types of folklore materials that I had recorded during my fieldwork in Ukraine should be processed for non-academic consumption. By choosing to transcribe and translate one type of audio file as opposed to another, the volunteers would reveal their interests while simultaneously doing some of the work needed for the publication of the texts. As my team and I worked with our contributors, we discovered that our site could also be used to glean information about the dynamics of heritage and ethnicity. Our volunteers came largely from the Ukrainian Diaspora, so we were crowdsourcing the audio files to a limited demographic rather than the public at large. We discovered that we could "experiment" with this group by adding select content to our site and watching volunteer reactions. Some of our results confirmed findings in other heritage situations while others were specific to Ukrainian culture and influenced by Ukraine's political history. The understanding of the Diaspora that we gained through work on this site could not have been possible through other means.

Ukrainian Folklore Audio (http://www.artsrn.ualberta.ca/folkaudio/)is a site where the public can listen to songs, stories, and beliefs recorded in Ukraine and among the Ukrainian Diaspora of Kazakhstan. Volunteers who wish to transcribe the recordings or translate them can "check out" the item that interests them and work on it. All completed transcriptions and translations are available for public use. Thus, when an item has been completely processed, the user can listen to a recording, see it written out in Ukrainian, and also view the parallel English translation. Anyone can use the site passively but, to avoid frivolous posts, we have required people submitting transcriptions and translations to register. Because of the registration requirement, we consider our site a modification of crowdsourcing and call what we do "groupsourcing." The registration requirement also helps with quality control. Submissions made by a volunteer are visible only to that volunteer and to the person monitoring the site (usually me) and are posted for general viewing only after they have been checked for accuracy.

Ukrainian Folklore Audio is an outgrowth of Ukrainian Folklore Sound Recordings (http://www.artsrn.ualberta.ca/UkraineAudio/). This is a research site developed with the help of Eric Zhang, a programmer then working at the Arts Resource Centre of the University of Alberta; Svitlana Kukharenko, a graduate student at the time; and Peter Holloway, a volunteer. The impetus behind the Ukrainian Folklore Sound Recordings site was my desire to easily find the information that I was seeking in the vast volume of recordings I had accumulated during my fieldwork. Like all folklorists, I faced the problem of dealing with a large volume of data, over 200 hours of sound in my case. Transcription of sound recordings—the standard way of processing field data—if done digitally, does produce files that can be searched. The problem with transcription is that it is an enormously time-consuming process and, once a sound file is converted into text, the expressive qualities of speech such as intonation, inflection, and volume are lost. Our solution was to index the sound files. We noted the time in a recording when a particular topic was being discussed. There were a limited number of topics since I was working on family rituals dealing with birth, marriage, and death. With the help of Zhang, we developed a program that linked each identified topic to the point in each sound file where the particular topic arose. Thus, if one clicks on a topic such as songs connected with weddings, the program will show a list of all sound recordings in which that topic is discussed. The researcher can then select a recording, click on it, and listen from the point where the discussion of the desired topic begins.

Figure 1 shows a screen capture of the opening page of the original site. In the lower portion are the major topics in both Ukrainian and English. These topics are in approximate chronological order. "The Wedding" has been marked with a red arrow. If "The Wedding" is clicked, the next web page shows the major sub-categories of "The Wedding." Each of these sub-categories, if clicked, leads to subsequent pages. As an example, we show in Figure 2 the
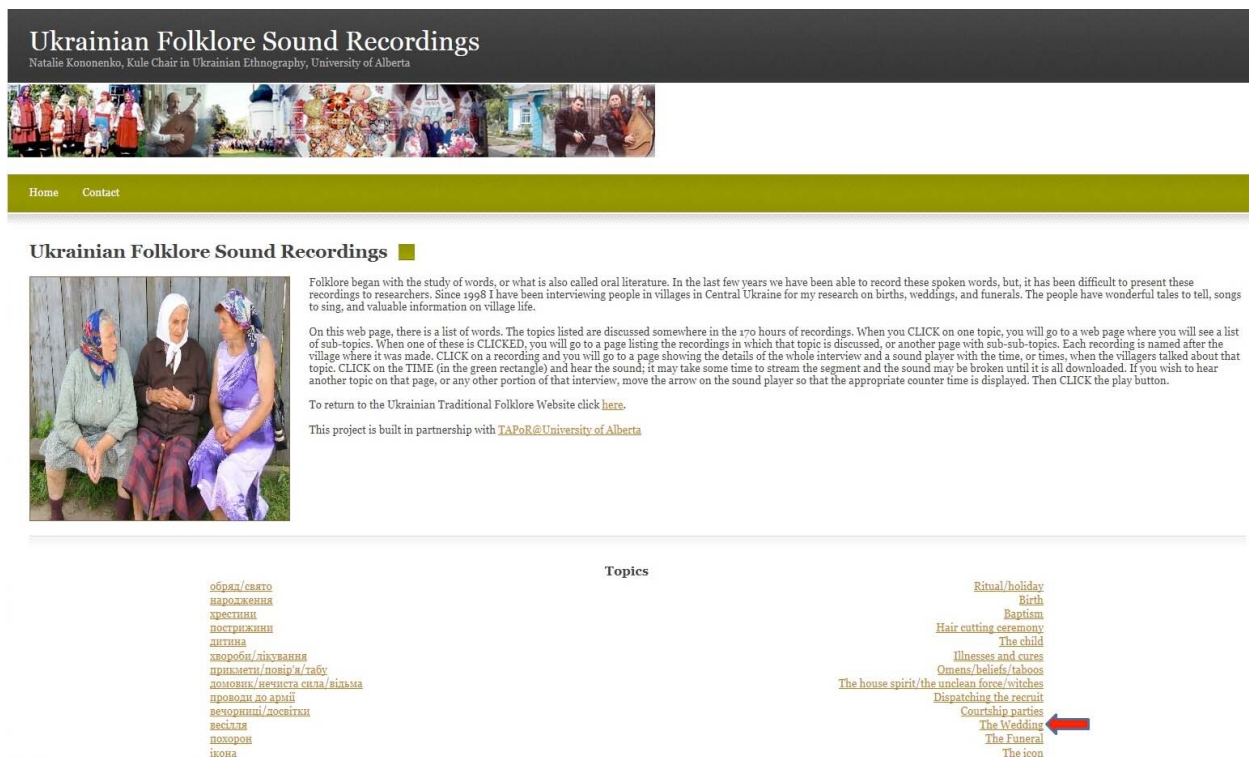

Fig. 1. Our original searchable audio website: http://www.artsrn.ualberta.ca/UkraineAudio/.

Fig. 2. An example of a search through "Wedding" subtopics.

sequential pages from a search of: The Wedding->Ritual songs->During the wedding->Singing of the clergy. The final screen capture in Figure 2 shows a partial list of recordings (named by village and year) where this topic was described in an interview. If one recording (iavorivka2000b) is clicked, the user sees the screen shown in Figure 3. This page gives detailed information on the interview and the arrow indicates that the topic in question was discussed between 06.35 and 15.28 minutes in the recording. Once the recording is downloaded, it will automatically play from the time indicated by the red arrow near the bottom of the page.

Fig. 3. An example of the background information to a sound file.

This website was a major breakthrough. It has enhanced my research tremendously because, in addition to taking me to all of the recordings where I can listen to information on topics such as wedding songs, it also allows me to listen to the context of my topic, should I chose to do so. Thus, I can move the cursor to a point before the time when the song in question begins and hear the conversation leading up to the singing, or I can listen to the part of the recording that follows the song and learn what was said after the song was finished.

While Ukrainian Folklore Sound Recordings is a boon to researchers, it is a research site, too big and cumbersome for use by the general public. The data found on this site are minimally processed, a situation that is beneficial for researchers because of the wealth of information included, but a hindrance to non-academic users. To make the material I collected in Ukraine accessible to the non-specialist, it needed to be trimmed down to select items, transcribed, and translated. But we were faced with two problems. One was again the problem of transcription and its time demands, a problem that would be compounded by adding translations. The other problem was selection. We could guess which items out of the vast amount of data I had collected would appeal to our target audience—not an ideal approach. Crowdsourcing presented itself as a good solution. The general public would help us with the time-consuming tasks of transcription and translation, and we would track the items chosen for processing to see which sorts of texts were of greatest interest to our contributors.

Crowdsourcing has been a very successful approach to engaging the public in research. The Sloan Digital Sky Survey used a crowdsourcing site called Galaxy Zoo (http://www.galaxyzoo.org/) and, within nine months, had 100,000 volunteers. Among them, the volunteers viewed and classified over one million galaxies, and each galaxy was viewed an average of 38 times (Clery 2011). Almost 200 volunteers contributed translations to *Suda* On Line (http://www.stoa.org/sol/), a project related to the tenth-century Byzantine encyclopedia. Our project was different from many other crowdsourcing efforts in that we were dealing with a heritage community. Emotions are a crucial part of heritage and our most interesting findings have to do with the role that heritage played in our project.

We built Ukrainian Folklore Audio in order to make it as user-friendly as possible. The interface, designed by Karl Anvik and graduate student Megan Sellmer, is simple in order to encourage contribution by users with all levels of familiarity with online work. For content, we decided to "cut out" discrete stories, songs, and accounts of magical beliefs and to give the texts only; information about the performer and the time and place of recording was not included. This decision to exclude performer information was based on the need for simplicity and on the desire to make the site as "familiar" as possible to potential non-academic contributors by following the "texts only" practice of print collections aimed at this demographic. Our hope was eventually to include pictures of and information about performers on the site if user interest and funding for technical support permitted. In the meantime we assumed that those people who did ask for performer information could be directed to the research site.

The "cutting" of sound clips was done by then graduate student Maryna Chernyavska. Deciding which types of texts to present was our most difficult task and was a cause for much debate. Three criteria determined our choices. One again involved print media precedents. Publications aimed at non-academic readers are typically collections of folktales or folk songs. This led us to decide on including stories and songs on our site. The second criterion was past

reaction both to the research site and to the presentation of data from my recordings. For instance, non-academic users searched the research site specifically for songs. For presentation to non-academic audiences, I do public outreach activities where I tell English versions of the tales told in Ukrainian on the research site. The story sessions are very popular. Audience members often asked for information on Ukrainian magic and these inquires led us to include short accounts of practices such as protecting the home from being struck by lightning. The third criterion was ease of transcription and translation. Songs, we thought, would be more difficult to transcribe and translate than prose and consequently a less popular choice among volunteers. This assumption led to the decision to offer two prose categories on the site and only one verse category.

Figure 4 shows the opening page of the Ukrainian Folklore Audio website. Anyone can listen to one of the 199 audio files and, if they have been processed, they can see the transcriptions and translations. Figure 5 shows some of the songs. If the first song is "viewed," the screen shown in Figure 6 appears. If the visitor has registered as a volunteer, the screen in Figure 7 is visible. This screen again allows the files to be heard and viewed, and if particular items are not "locked," they can be "signed out" for the volunteer to work on transcribing or translating.

When Ukrainian Folklore Audio went online in 2011, I sent announcements to various newsgroups, both general Slavic and specifically Ukrainian. The announcements were greeted with enthusiasm, especially by members of the Ukrainian groups. Unfortunately, this enthusiasm did not translate into a plenitude of volunteers. Many people used the site passively and recommended it to others, but did not contribute. We sought to discover the reasons for lack of participation and
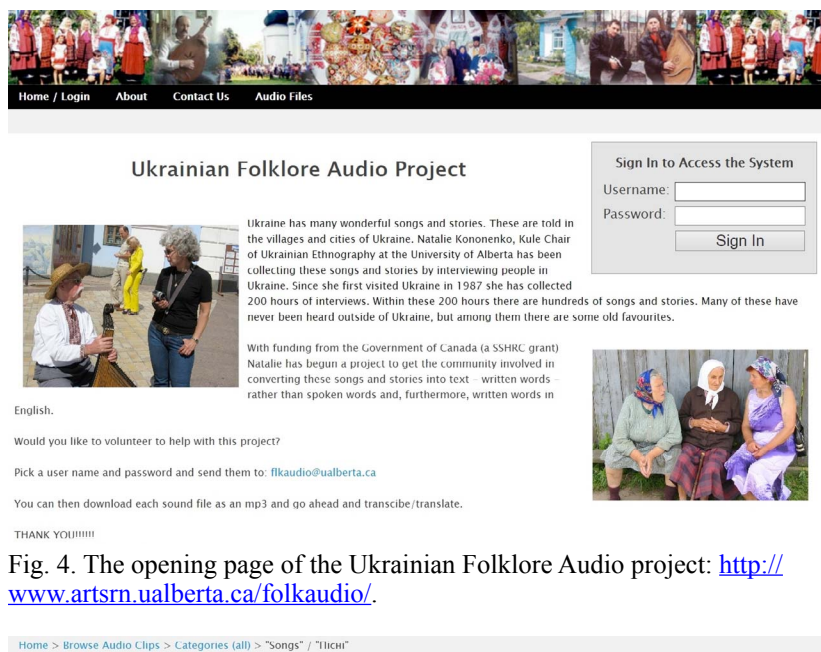

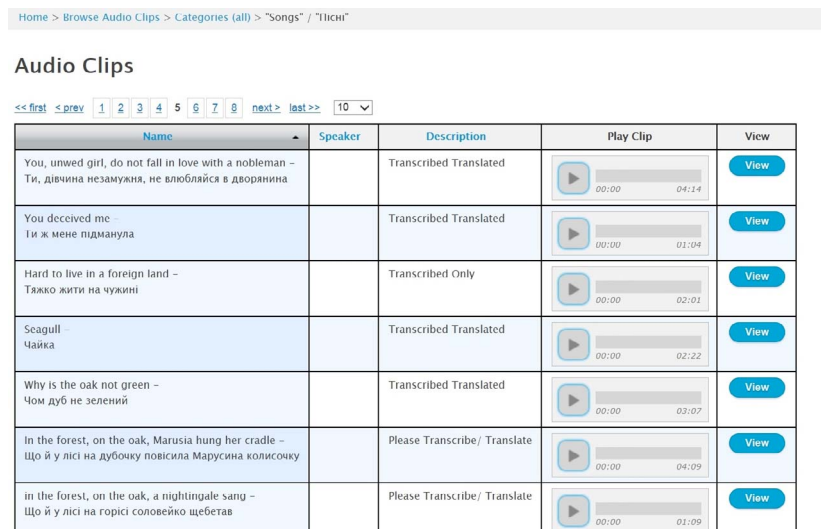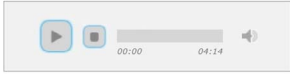Fig. 4. The opening page of the Ukrainian Folklore Audio project: http://www.artsrn.ualberta.ca/folkaudio/.

Fig. 5. An example of a page showing already processed files and files available for processing.

Fig. 6: An example of "published" transcriptions and translations.

Fig. 7. An example of the page that a registered visitor sees.

considered offering incentives. As pointed out by Hars and Ou (2000), there are various ways of encouraging people to contribute to online work, the most common being "gamifying" the site and turning the tasks into a contest. We decided against this approach because members of the Ukrainian Diaspora are quite insistent that all things Ukrainian must be treated seriously. We feared that introducing leaderboards and other elements of play would "trivialize" the project and offend cultural sensibilities. As we observed work on the site and interacted with newsgroups and individuals, it became clear that our problem was not so much incentives as disincentives. One issue was language. Many members of the Ukrainian Diaspora feel insecure about their language knowledge, and potential contributors feared criticism for posting transcriptions that were not in correct Ukrainian, a fact that several correspondents admitted to me in confidence.

For many heritage groups language is a mark of identity. For Ukrainians, this is doubly so. Ukrainians who emigrated to Canada were under pressure to switch to English and, since the first wave of emigration came in the late 1800s, that pressure was quite brutal, with children being beaten in school for speaking Ukrainian and forced to adopt English names. Attacks on the Ukrainian language were very much a part of life in Ukraine as well, both when this country was part of the Russian Empire and subsequently when it became part of the Soviet Union. Thus, being able to speak Ukrainian was a mark of resistance and highly valued by anyone with nationalist feelings. The people who were interested in things Ukrainian—the potential contributors to our site—were the ones who prized language retention. They also valued a pure Ukrainian, one not contaminated by forced Russification or full of Anglicisms. Unfortunately, the hyper-correct Ukrainian that was considered politically correct was not something that most

people could master. Paradoxically, language instruction in the schools—for which Ukrainian organizations in Canada fought long and hard—worked against language retention. By revealing the differences between grammatically correct Ukrainian and the language spoken in the home, formal instruction discouraged language use. An identical phenomenon was observed by Rocky Sexton (2000) in French Louisiana, where formal language instruction by teachers brought from France discouraged language use by revealing the differences between the French spoken in France—a version considered more prestigious and correct—and the language used by French speakers in Louisiana. What this situation produced in our case was potential contributors who wanted to retain the language but felt their own Ukrainian was not up to the ideal that they sought to promote. We anticipated this problem and made it clear that no transcription or translation would be made public until it was first checked by a native speaker. We also offered the option of anonymity to people who wanted to participate in the project but feared criticism.

Another language issue had a different emotional resonance. Ukrainians started coming to Canada toward the end of the nineteenth century, but the Soviet Socialist Revolution closed the borders of Ukraine and Ukrainians in Canada had only limited contact with their homeland. During the seventy years that the Soviet Union existed, the Ukrainian spoken in Canada and the language spoken in Ukraine itself developed independently and grew apart. This linguistic evolution made modern Ukrainian, the language in my sound recordings, somewhat difficult for Ukrainian Canadians to understand. The Diaspora interpretation of the differences between their Ukrainian and that spoken in their homeland was that the latter must be wrong. Members of the Diaspora felt that they had preserved the pure form of the language while the spoken language in Ukraine had been corrupted by forced Russification. Our supposition that people would be eager for material from the country they had been unable to access for so long was thus thwarted by the Diaspora's feeling that the language and culture of Ukraine had been corrupted by Soviet rule and was not genuinely Ukrainian. The language of the sound files was "wrong" and people were reluctant to work with it.

The type of text chosen for transcription and translation was unexpected. For the first year or so of the site's existence, songs and only songs were transcribed and translated. In fact, we added more songs to meet demand. Examination of the preference for songs quickly explained this phenomenon. As Halpern and Barlett (2011) have noted, a song stays in the mind as something called an "earworm"—sometimes against a person's will. Thus, songs are better remembered than ordinary speech and may often retain older linguistic forms, in this case making them more familiar to Ukrainian Canadians. Music surrounds the listener and elicits a strong emotional response, thus making it better suited to the nostalgia and other sentiments that are part of heritage. Any perceived irregularities—whether in syntax, in word choice, or in pronunciation—are less apparent in music where they can be explained as distortions made to fit verse structure, thus obviating the language issues discussed above. People also gravitated toward songs because they were familiar, something that the contributors had heard before. We expected that people living in the West, Diaspora Ukrainians included, would be attracted to the new. This is not the case when it comes to heritage, and we repeatedly encountered the concept of "preservation of the old" in all of our heritage work.

The crowd- or groupsourcing approach allowed us to test the observations above. The fact that potential contributors did not work with prose texts led us to hypothesize that they were

deterred by language. Many of the attractive features of songs are also found in at least one other category on our site: folktales.[1] Like songs, folktales are associated with heritage and are well-known to the Diaspora through print media. But while the songs on our site have words similar to those in published collections, the tales do not. The oral style of songs is retained in print in order to keep the rhythm and the rhyme. This is not true of tales where colloquial language is rendered in grammatically correct Ukrainian for publication. The tales on our sound files are, of course, oral and are told in a colloquial language that differs from the literary standard. Furthermore, the tales are in the language spoken in contemporary Ukraine, not the language of the Diaspora, leading potential contributors to suspect Russification. If the language issue were removed, we proposed, the tales would become as attractive as we had originally assumed. To test this theory, we transcribed some of the texts ourselves. Graduate student Myroslava Uniat wrote out the tales, giving the words as they were spoken in the sound file and also providing in parentheses the literary standard version of any dialectal or colloquial words. Sure enough, once the tales were transcribed, a volunteer from North America began translating them. The fact that the colloquial language of the prose texts hindered transcription was also confirmed by another event. A volunteer who had done a number of the tale translations submitted a translation of one of the magical belief texts. She did not transcribe the text, perhaps deterred by the colloquial speech and worried that her Ukrainian was not up to the task of supplying both the spoken and the grammatically correct forms, but she did understand the sound file and was able to render it in English.

To test if familiarity with a song determined its attractiveness to volunteers, we posted a set of songs recorded from Ukrainians living in Kazakhstan. Our premise was that these songs would be less familiar and therefore less attractive. These songs have been transcribed by a person living in Ukraine, not by a member of the Ukrainian Diaspora. This result does indicate that our hypothesis was correct. Ukrainians in Ukraine live in greater geographical proximity to ex-patriots living in Kazakhstan; also, the citizens of both countries share a Communist past. Their expressive cultures are therefore more likely to be similar and the songs of one group are more like those of the other. To date, no Diaspora Ukrainians have worked with the Kazakhstan songs, either as transcribers or as translators.

The aforementioned observations were presented at conferences, and the feedback we received, combined with observations of other crowdsourcing websites, suggested that frequent, if not constant, interaction with volunteers stimulates contributions. This feedback prompted posting news about the site to various newsgroups. Every few months an update on new contributions is sent to social media These updates partially replace things like leaderboards and other incentives by providing public recognition of contributors' work. Every posting brings in a few new contributors and stimulates increased activity among those already working on the site.

As for the future of the site, we plan a dual approach. As we see what material appeals to the public, we will post more of it, thus doing what we had originally intended and giving the non-academic users the type of items that they want. We also plan to continue to experiment. We originally posted only short items on the supposition that short texts would attract volunteers. Now that several volunteers have told us that they work on the site for relaxation, we will try

---

[1] Two examples of Ukrainian folktales in English translation are Bloch 1964 and Zhelenova 1986.

posting longer, more complex, and thus potentially more appealing texts. Public reaction to long folktales should be interesting. Long tales are particularly engaging because of their complex plots and their artistry. They envelop the listener with enchantment much as songs do with music. Will people find these tales attractive enough to work on despite their length? Only time will tell.

In sum, Ukrainian Folklore Audio is a crowdsourcing site that permits the researcher and the general public to have a dialogue without words. Instead of asking our public questions directly, we post materials on our site. User reactions to the materials that we post, namely their choices of which materials to transcribe or translate, is perhaps a truer response to our queries than any direct question. The posting of materials and the availability of choice gives us a better sense of Diaspora language issues. Because language retention (or its lack) carries such emotional weight, users reveal through their behavior positions that they would probably not admit to if asked directly. Similarly, heritage issues are loaded with emotion, and approaching them through a folklore website that allows user choice gives us information about attitudes toward heritage that we would not have been able to get through direct questioning.

University of Alberta

References

Blank and Howard 2013   Trevor J. Blank and Robert Glenn Howard, eds. *Tradition in the Twenty-First Century: Locating the Role of the Past in the Present*. Logan: Utah State University Press.

Bloch 1964   Marie Halun Bloch. *Ukrainian Folk Tales*. New York: Coward-McCann Publishers.

Clery 2011   Daniel Clery. "Galaxy Zoo Volunteers Share Pain and Glory of Research." *Science*, 333.6039:173-75.

Halpern and Bartlett 2011   Andrea R. Halpern and James C. Bartlett. "The Persistence of Musical Memories: A Descriptive Study of Earworms.*" Music Perception*, 28.4:425-31.

Hars and Ou 2000   Alexander Hars and Shaosong Ou. "Working for Free? Motivations for Participating in Open-Source Projects." *International Journal of Electronic Commerce*, 6:25-39.

Kononenko 2013   Natalie Kononenko. "Groupsourcing Folklore Sound Files: Involving the Community in Research." *Canadian Slavonic Papers/Revue canadienne des slavists*, 55.1-2:131-50.

Marett and Barwick 2003   Allan Marett and Linda Barwick. "Endangered Songs and Endangered Languages." In *Maintaining the Links: Language, Identity and the Land.*

Proceedings of the Seventh FEL Conference, Broome, Western Australia. Ed. by R. McKenna Brown and J. Blythe. pp. 144-51.

Ong 1982   Walter J. Ong. *Orality and Literacy: The Technologizing of the Word*. London and New York: Routledge.

Sexton 2000   Rocky L. Sexton. "Cajun-French Language Maintenance and Shift: A Southwest Louisiana Case Study to 1970." *Journal of American Ethnic History*, 19.4:24-48.

Zhelenova 1986   Irina Zhelenova, trans. *Ukrainian Folk Tales*. Kyiv: Dnipro Publishers.