

# **Princesses in Grimms' folktales and Andersen's literary fairy tales:a corpus study**

Yingxin Gao  
University of Wisconsin-Madison

## **1 Introduction**

When it comes to fairy tales, the first name popping up in our mind is probably Brothers Grimm. Indeed, it is ranked as the second most widely read book across the world, following *Bible*. Many of us grow up listening to Brothers Grimm's bed-time stories. And the key word here is 'listen', which triggers the idea of this study, a register analysis of Brothers Grimm's folktales in comparison with the spoken register, conversation. What may surprise most of us is the fact that, world-renowned as children stories though, the Grimms' folktale collection actually began as a scholarly project aiming at the adult audience.

In the eighteenth century, French stories were popular in Germany and later Germany was under Napoleon's rule. In the hope of seeking consolation and refreshment in German literature and language as well as fostering a national identity and unification, Jacob Ludwig Grimm(1785-1863) and Wilhelm Carl Grimm (1786 - 1859), better known as Brothers Grimm, began to collect German folktales at the beginning of the nineteenth century. However, instead of going on to the field and recording stories told by the peasants, as most folkloric scholars do today, they mainly invited to their household story-tellers, most of whom were educated young women from the middle class or aristocracy; the guests told tales aloud while Brothers Grimm

noted down on first or after several hearings (Zipes, 1999). The Brothers' primary concern was to "uncover the etymological and linguistic truths that bound the German people together" (Zipes, 2002, p.25). Yet, it's still under dispute to what extent these stories are of genuine German origin since quite a lot of the stories the Grimms' guests told about were heard from their French governesses, nursemaids or servants (Zipes, 1999). Some stories from their collection remind us of the seventeenth-century writer Charles Perrault's famous folktale collection *Mother Goose Tales*. To name a few, the plots and characters of the Grimms' *Little Red Cap*, *Puss in Boots*, *Briar Rose*, *Cinderella*, and *Bluebeard* are very similar to Perrault's *Little Red Riding Hood*, *Puss in Boots*, *The Sleeping Beauty in the Wood*, *Cinderella: or, The Little Glass Slipper*, and *Bluebeard*.

The first volume of Brothers Grimm's folktales, *Children's and Household Tales* came out in the year of 1812, but was criticized by the public for its weird content and its particularly plain and unadorned language due to faithfulness to the true oral tales. They sold badly. When preparing for the second volume, Brothers Grimm targeted their book at children and their parents from the middle class. Intent on genuine and faithful record of folktales though, this time, they edited to embellish the language, elaborated on the plot, enriched characters, added stylistic features as well as censored proper content for some pedagogical purpose. The second volume of *Children's and Household Tales*, published in 1815, sold extremely well and established Brothers Grimm fame and fortune. After that, they published new volumes several times until 1857 when the seventh edition came into being and was settled as the final version, which has been circulated around the world to date. Therefore, even though originated from oral tradition, their collection was undergone editing and adapting. Zipes (2002) compares the same stories (*Snow White*, *Rapunzel*, etc.) in the 1857 edition and the 1812 edition and finds that stories the later edition are longer than the first one. What's more, from the very beginning, most of their story-tellers were educated; their narration, despite in the spoken form, might be influenced by literary sources (Tatar, 2004). Zipes (1999) points out that "the Grimms took tales directly from books and journals and edited them according to their taste" (p.70). Hence, the Grimms' collection is probably situated in a paradoxical status. Is it more appropriate to claim that it belongs to the spoken register in the form of transcription or the written register? Some scholars place more emphasis on Brothers Grimm's editing execution and treat their folktales as written stories while other scholars prefer to regard these folktales as oral tales. For example, Tatar (2004) argues that "they reached back to an oral storytelling tradition that is fundamentally different from literary culture"

(p.xxxiv). In this paper, I will explore to what extent the Grimms' folktale collection follows the original oral narration based on a register analysis on its closeness to the spoken register conversation compared with Hans Christian Andersen's literary fairy-tales.

Born into a poor family, Hans Christian Andersen was ashamed of his lower class family background so he desired for a social rise through writing. He started out as a dramatist but it did not work out. Later he wrote fairy tales which catered to a growing audience of middle class children and their parents as well. Bearing in mind the dream of a social rise, he aspired to write fairy tales embedded with the bourgeois essentialist ideology in the hope of appealing to the upper class. It should not be surprising to find the Christian notion of God, sacrifice, suffering and transcendentalism widely scattered in his stories such as *The Little Mermaid*. This time, he established himself as a reputed writer.

Similarly, Brothers Grimm's collection also touches on topics like Christianity, Protestant work ethic, social justice, patriarchy, etc. Among both collections, princess stories are very outstanding. I will also focus on linguistic features addressing the princesses so as to figure out differences among the princess image in the folkloric tradition and the literary tradition.

## **2 Methodology**

In this study, I created my own corpus and manually counted the pervasive linguistic features throughout the text. The small corpus was about twenty thousand words comprised of five stories from Brothers Grimm's collection - *Rapunzel*, *Little Snow-White*, *The Twelve Dancing Princesses*, *Briar Rose*, *Cinderella*, and five tales from Andersen's collection - *The Princess and the Pea*, *The Wild Swans*, *The Snow Queen*, *The Marsh King's Daughter*, *The Little Mermaid*. To strike a balance among the text lengths of both, I included full texts of Brothers Grimm's folktales as they are short and random excerpts of Andersen's lengthy literary tales. To assemble a random text for Andersen's collection, I picked out the third paragraph of each page until it reached around two thousand words for each story. If the excerpt of all third paragraphs was still too short, I would add the first paragraph of each page until it fulfilled the length requirement. Nonetheless, the full text of *The Princess and the Pea* was included since it was less than four hundred words. Eventually, the sample text of the Grimm's folktales totaled 10070 words while that of Andersen's literary fairy-tales amounted to 9085 words.

Biber, Conrad and Leech (2002) claim that conversation demonstrates

high frequencies of pronouns and verbs, low type-token ratios, and simple noun phrases while the written register presents more noun phrases, higher type-token ratios, and complex noun phrases. So I counted the occurrences of pronouns, noun phrases (NPs) excluding pronouns, subject and object positions of noun phrases, verb phrases (VPs), and adjective phrases (AdjPs), calculated the type-token ratio of noun phrases and examined the complexity of noun phrases in terms of types of modifiers like adjectives and prepositional phrases.

To address the personalities of princesses in the folklore and literary fairy tales, I counted occurrences of different verb phrases clustered into different semantic groups. I first divided them into passive voice and active voice and then categorized the active verb phrases into dynamic activity verb phrases like 'run away', static activity verb phrases like 'sleep' and 'lie', communication verb phrases like 'say' and 'shout', mental verb phrases like 'think' and 'believe', and existence verb phrases like 'is pretty' and 'has a dog'. Of all these linguistic features, high frequencies of subject position of noun phrases and dynamic activity verb phrases signal an active personality of the princess.

### **3 Data Analysis**

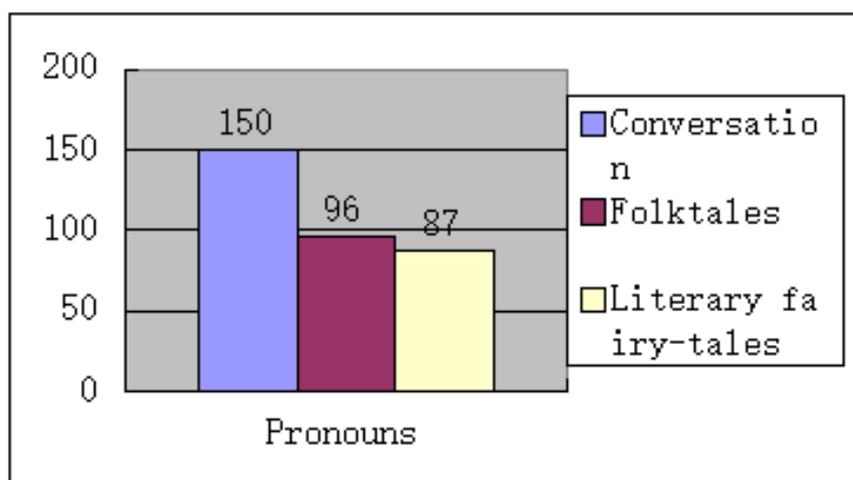
#### **3.1 Are Brothers Grimm's folktales closer to conversation than Andersen's fairy tales?**

To address this research question, I scanned through the texts and counted frequencies of linguistic features related to conversation-pronouns and verb phrases, I figured out the type-token ratio of noun phrases referring to princesses and I investigated the complexity of noun phrases.

First, I compared the frequency of pronouns in Brothers Grimm's folktales and Andersen's literary fairy tales and conversation.

study

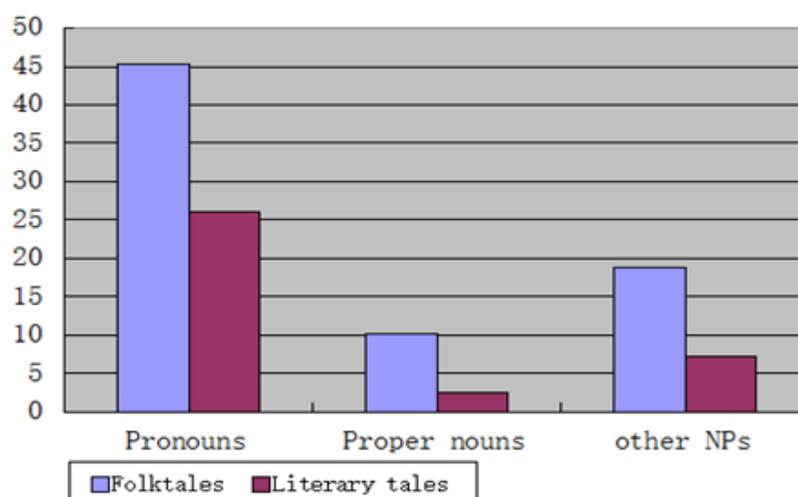
Figure 1 Normed rates of occurrences of pronouns (per 1000 words) in conversation<sup>1</sup>, Brothers Grimm's folktales and Andersen's fairy tales



From Figure 1, it is clear that the occurrence of pronouns in the Grimms' tales is just about two thirds of that in conversation and is only slightly higher than that in Andersen's fairy tales.

However, when narrowing down to pronouns referring to princesses, the gap between the Grimms' folktales and Andersen's fairy tales is much wider, as shown in Figure 2 and Figure 3.

Figure 2 Normed rates of occurrences of NPs referring to princesses in Brothers Grimm's folktales and Andersen's literary fairy tales



<sup>1</sup> The number for conversation is based on Figure 4.7 in *Longman student grammar of spoken and written English* (Biber, Conrad & Leech, 2002, p. 93).

Figure 3 Type-token ratios of princess NPs in Brothers Grimm's folktales and Andersen's fairy tales

	Folktales	Literary fairy tales
Type	27	24
Token	501	304
Type-token ratio	0.05	0.08

From Figure 2, we can see that per thousand words, pronouns occur 45 times in the Grimms' folktales, almost twice as the occurrence of pronouns in Andersen's literary fairy tales (25 times). Comparing Figure 1 and Figure 2, we can find that pronouns referring to princesses take almost half of all pronouns in the Grimms' folktales while those in Andersen's literary tales account for less than one third of all pronouns. The cause behind this is probably due to different production circumstances of the original creation of folktales and literary tales. In the story narrating at Brothers Grimm's household, story-tellers produced stories in real time so they did not have much time to think of different noun phrases to refer to the princess. Besides, they needed to clearly tell the audience who did what and the easiest way is using proper nouns, accounting for the higher frequency of proper nouns in the folktales. This also results in a lower type-token ratio of noun phrases in the Grimms' tales. On the contrast, Andersen could spend as much time as he could to come up with all different noun phrases referring to the same character. Furthermore, he could add a lot of details and modify the noun phrases with all kinds of modifiers to embellish his language. Hence, the Grimms' text demonstrates a lower type-token ratio than Andersen's text.

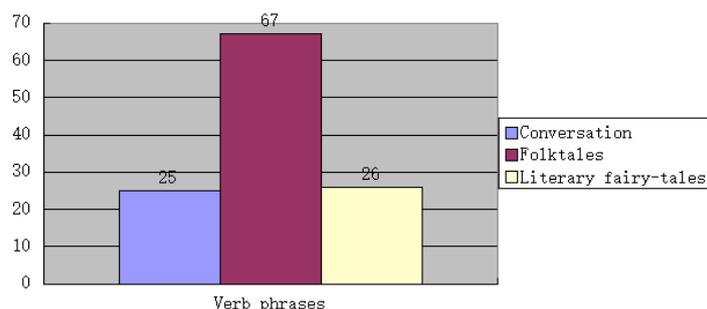
Another linguistic features of the spoken register or conversation is a high frequency of verb phrases.

Figure 4 Normed rates of occurrences (per 1000 words) of VPs in conversation<sup>2</sup>, Brothers Grimm's folktales and Andersen's literary fairy tales

---

<sup>2</sup> The number for conversation is based on Figure 2.1 in *Longman student grammar of spoken and written English* (Biber, Conrad & Leech 2002: 23).

*study*



Surprisingly, Figure 4 shows that both the Grimms' and Andersen's tales outnumber conversation in the frequency of verb phrases. The Grimms' tales almost triple the conversation register. Their different communicative purposes contribute to this result. Apart from the informational purpose of telling people what happened, conversation also carries the purpose of strengthening personal relationship or social bonds so there are some verb phrases but not too many. Nonetheless, the main purposes of both Brothers Grimm' and Andersen's tales are to narrate a story, which requires a lot of verb phrases. Again, the big gap between the frequencies of verb phrases in the Grimms' and Andersen's stories is attributed to the production circumstances. In real time production of the oral folktales, the story-tellers were not able to add a lot of modifiers but simply presented to the listeners all kinds of actions or activities. Andersen, nevertheless, was not only able to modify the verb phrases but also add lots of description about the environment and so on, which in turns lowered the percentage of verb phrases per thousand words.

In short, communicative purposes and production circumstances contribute most to the differences of linguistic features among conversation, folktales and literary fairy-tales. In comparison with Andersen's literary fairy-tales, Brothers Grimm's folktales present a higher frequency of pronouns and verb phrases but a lower type-token ratio, signifying a closer relationship with the spoken register. However, in the face of actual conversation, the frequency of pronoun in the Grimm's stories is much lower. That is to say, the Grimms' tales are in a mediate position between the spoken and written register.

### **3.2 Are princesses in Grimms' folktales more active than their counterparts in Andersen's literary fairy-tales?**

In order to scrutinize at whether princesses in Andersen's fairy-tales are more passive than those in the Grimms' folktales, the author counts occurrences of the subject position and object position of princesses, active verb phrases and passive verb phrases as well as different types of verb phrases according to

semantic classification.

Figure 5 Normed rates of occurrences (per 1000 words) of subject and object positions of princess in Brothers Grimm's folktales and Andersen's literary fairy tales

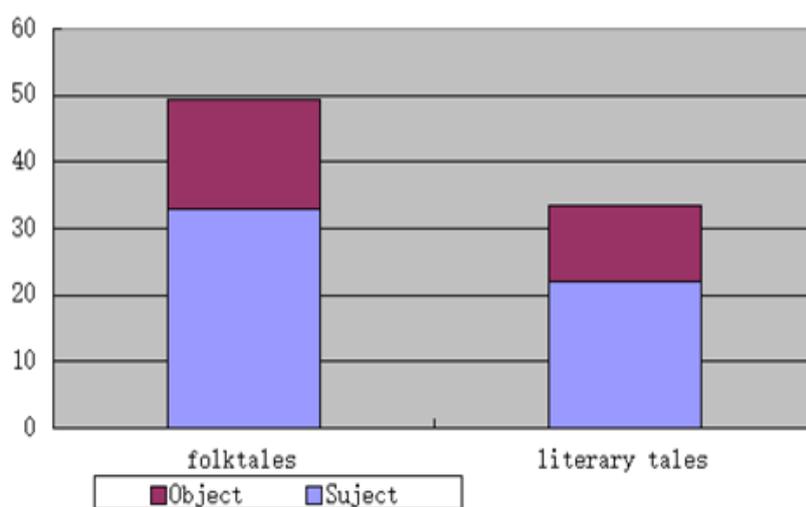


Figure 5 shows that the overall occurrences of princesses in the Grimms' folklore are higher than in Andersen's literary fairy tales, indicating a more significant role of the princess in folktales than in literary fairy tales. However, the percentage of subject position of princess noun phrases in folktales is lower than that in literary fairy tales, which, in contrast, suggests that the literary princesses are more active than their folkloric sisters. But is this the case? Does a subject position always signify activeness of the princess?

I examined all verb phrases with the princess as the agent and grouped them into five semantic categories: dynamic activity verb phrases ('run', 'go', etc.), static activity verb phrases ('sleep', 'lie', etc.), communication verb phrases ('say', 'talk', etc.), mental verb phrases ('think', 'want', etc.) and existence phrases (copula 'be', 'have', etc.). Obviously, in statements like 'she is beautiful', 'she has a fancy dress', 'she wants to go to the ball', 'she sleeps for a hundred years', the princess is actually doing nothing even though she is in a subject position. Therefore, among the five verb groups, only a high frequency of dynamic activity verb phrases or communication verb phrases can be accounted for the activeness of princesses.

*study*

Figure 6 Occurrences of different verb phrases in Brothers Grimm's folktales and Andersen's literary fairy tales

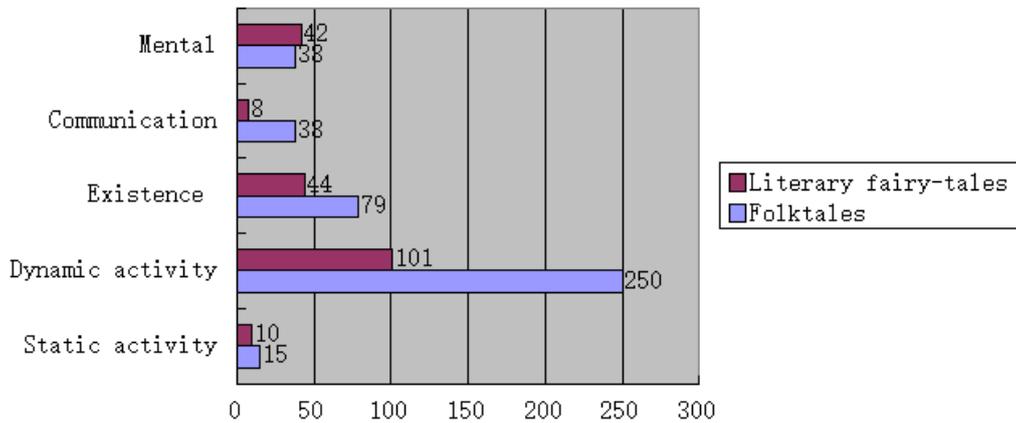
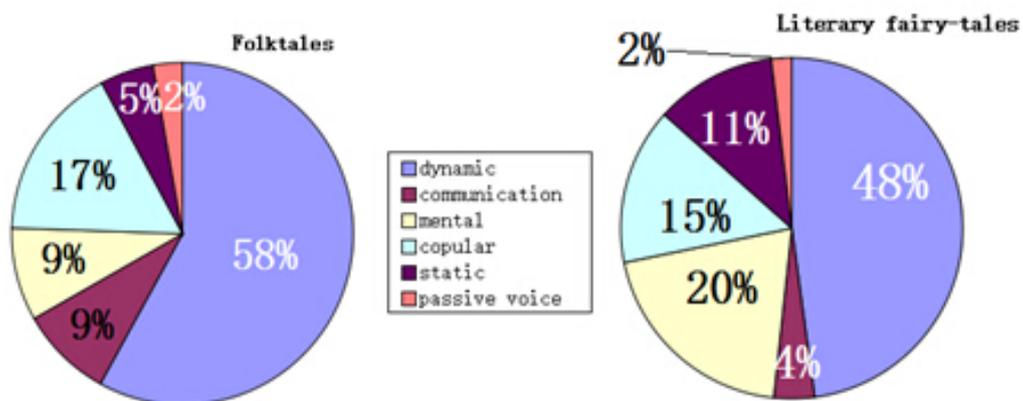


Figure 6 clearly shows that our folkloric princesses outperform our literary princesses in all activities except for mental activities. However, while folkloric heroines ‘run’ more frequently and ‘think’ less frequently, they ‘lie’ motionlessly more often too, which cancels out the activeness of the folkloric princesses. To further explore this, I took a look at the percentages of all verb phrases within the folkloric arena and literary arena separately.

Figure 7 Percentages of all verb phrases referring to princesses in Brothers Grimm's folktales and Andersen's literary fairy tales



Based on Figure 7, it seems that folkloric princesses are not much more active than their literary sisters, with slightly over half of the action being dynamic verb phrases in Brothers Grimm's folktales and slightly less than half

*study*

of the action being dynamic verb phrases in Andersen's literary fairy tales. The percentages of copula verb phrases between the two are almost the same while both have only 2% of verb phrases occurring in a passive voice. The most obvious gap lies in mental verb phrases, which accounts only 9% of all verb phrases in the Grimms' folklore but 20% in Andersen's fairy tales. The literary princesses obviously think more than their folkloric counterparts. However, the overall difference between folkloric and literary princesses' activeness is not very significant. Perhaps the princesses in folktales and literary folktales are after all quite similar. And a detailed examination at the description of princesses and their antagonists, usually female, confirms this prediction.

Figure 8 Most common Adjective Phrases modifying Princess and Female Antagonist in Brothers Grimm's folktales and Andersen's literary fairy tales

	Folktales	Tokens	Literary fairy tales	Tokens
Princess	beautiful	8	beautiful	3
	young	7	real	3
	little	6	little	9
	dear	6		
Female Antagonist	old	14	deceitful	1
	wicked	4	wicked	1
	false	3		

Figure 9 Most common NPs referring to Princess and Female Antagonist in Brothers Grimm's folktales and Andersen's literary fairy tales

Princess			
Folktales	Tokens	Literary fairy tales	Tokens
princess	30	princess	27
child	14	child	7
girl	13	girl	2
daughter	9	daughter	4
sister	3	sister	6
wife	3	wife	2
Female Antagonist			
Folktales	Tokens	Literary fairy tales	Tokens

*study*

queen	18	queen	1
woman	18	other	3
enchanted	15	witch	2
stepmother	11		
sister	7		

Figure 8 shows that the most common adjectives depicting princesses in both Brothers Grimm's and Andersen's tales are 'little' and 'beautiful'; however, the female antagonists in both collections are described as 'bad', 'old' and 'wicked', if not referred to in bare noun phrases. As for the noun phrases addressing the princesses, the most frequent nouns in both texts are 'princess', 'child', 'girl', 'daughter', 'sister' and 'wife' while nouns like 'woman', 'enchanted', 'queen' and 'stepmother' are most often associated with female antagonists (Figure 9). This may reflect the nineteenth century's female stereotypes: the ideal female should be delicate, innocent and their main roles are the obedient daughters and wives; the sophisticated and mature 'woman' and 'queen' are labeled as 'wicked' 'enchanted', mirroring the male's inner fear of female with knowledge.

#### **4 Conclusion**

The way Brothers Grimm collected their oral tales by inviting guests, most of whom were young middle or upper class ladies, and their later editing practice place their text in a mediate position between the spoken register and written register. Similar to conversation, the real-time production circumstances of story-telling resulted in a high frequency of pronouns and proper nouns as well as a relatively low type-token ratio concerning the princesses. The shared temporal context and physical space of story-tellers and listeners led to a relatively high occurrence of pronouns but not so high as in conversation, since the story-teller was trying to narrate a story usually involving many characters. To narrate the story in a clear way, the speaker used proper nouns very often since pronouns could be confusing in this case. Also, because of the narrating purpose, the narrator adopted lots of verb phrases to present action in the story, much more than in the conversation whose primary communicative purpose is to establish social bonds instead of describing action. Andersen's fairy tales were probably revised and edited many times in a carefully planned manner. Naturally, his text demonstrates the lowest frequency of pronouns

among the three and presents a higher type-token ratio than the Grimms' text. Due to the narrating purpose, Andersen's fairy tales are also embedded with many verb phrases.

Despite differences in linguistic features, both Brothers Grimm's and Andersen's stories have established a similar princess image: a fragile, pretty, little girl stirring male's protecting desire. Brothers Grimm's censorship and Andersen's intentional screening of proper contents in the hope of appealing to the traditional values confine their female heroines to the ideal domestic prototypes of 'daughter' and 'wife' in the nineteenth century. In contrast, both collections present the antagonist females as 'wicked' and 'old', reflecting male's hidden anxiety over female's maturity and potential capabilities of intricate plotting. The similarities in terms of the ideal female image in the nineteenth century, again, confirms Brothers Grimm's censorship and editing effort and proves that their collection is no longer genuine oral folklore.

## **References**

- Biber, Douglas, Susan Conrad & Geoffrey Leech. 2002. *Longman student grammar of spoken and written English*. Essex: Pearson Education.
- Tatar, Maria. 2004. Reading the Grimms' children's stories and household tales: origins and cultural effects of the collection. In *The Annotated Brothers Grimm*, ed. Maria Tatar, xxviii - xlvi. New York and London: W.W. Norton & Company.
- Zipes, Jack David. 1999. *When dreams came true: classical fairy tales and their tradition*. New York and London: Routledge.
- . 2002. *The Brothers Grimm: from enchanted forests to the modern world (2<sup>nd</sup> edition)*. New York: Palgrave Macmillan.

University of Wisconsin-Madison  
Department of English  
600 North Park Street,  
Madison, WI 53706

gao67@wisc.edu