

CORPUS LINGUISTICS AND CREATIVITY: STORYTELLING FOR UNIVERSITY STUDENTS

Ian Robinson

(University of Calabria, Italy)

Abstract

This article reports on an experience of using corpus linguistics to aid students in writing a creative text and so enhance their motivation. The work investigates the available literature to help understand what is meant by 'creativity'. Although used a lot in EFL, this word has complex meanings connected to it and there is a call to discuss it in a wider context. A worksheet was prepared for students using a corpus linguistic analysis of modern, English versions of the stories of the Brothers Grimm. The article describes how this worksheet was constructed with the use of a specialized corpus and the creation of a stop-list. This worksheet contained single words as well as word clusters found within the tales. Social Work students in an Italian university, involved in an English for Specific Purposes course, were then asked to use a selection of these words and phrases to help them write stories. The final stories were analysed and a follow-up questionnaire was used to elicit the students' perceptions concerning creativity. The study is linked to Dörnyei's work on the Motivational Self and tries to show how creativity in EFL can be seen as a motivational tool. The work concludes that creativity is essential in EFL and that it is something that can and must be fostered in students. It aims to help improve understanding of what we mean by creativity and to encourage students and teachers to adopt corpus linguistics within their own specific field of English learning and teaching.

*Creativity is not an accident, not something that is genetically determined.
It is not a result of some easily learned magic trick or secret,
but a consequence of your intention to be creative
and your determination to learn and use creative-thinking strategies.*

Michalko 2006, xvii

1. Introduction

There is a growing concern in the EFL world with creativity: this is seen in publications as well as conferences. However, very often the word is used without looking at its meaning and without trying to understand if this is applicable, or useful, in classroom settings. This article proposes an initiative aimed at encouraging undergraduate creativity in the use of English as a second (L2) or foreign language (FL) through the creative writing of a distinctly recognizable genre, which in the experience reported here is that of a fairy tale, with the aid of corpus linguistics. Successfully completing a creative task like this fits into Dörnyei and Ushioda's (2009: 29) L2 Motivational Self

System, in which they posit the three components of the system as *Ideal L2 self*, *Ought to L2 self* and *L2 learning experience*. A positive learning experience fits into the third category as an aid to motivation.

Students who use English as an L2 might not be aware of the language expected to be used in writing a specific genre. Being successfully creative in the L2 should act as a motivating factor for further studies and use of that language. I suggest that by applying Corpus Linguistic analysis to a specific genre of writing, in this case the stories of the Brothers Grimm, linguistic features of that genre would appear and could be foregrounded and act as an aid to being creative. Corpus linguistics has a long history in the literature of its pedagogical uses (see for example McEnery *et al.* 2006 and O’Keeffe, McCarthy and Carter 2007).

The students involved in this project were from a degree course in Social Work. It might seem strange to connect the world of social work with that of fairy tales. However, Gring-Pemble (2003: 9) writes, “The narrative paradigm holds that humans are essentially storytellers who create and communicate stories that form understanding, guide collective reasoning, and shape behaviour”. This narrative paradigm is linked to social work and social services by Slessor (2014: 422) when she states that “It is society that is [...] disabling, as opposed to the impairments of the individual, the social causes of disability can be found in a number of cultural practices, social attitudes and approaches [...] which seem to offer a stereotype of disabled people”. As an example of where we can find this happening in culture she says (*ibid.*), “In fairy tales, in particular the tales associated with the Brothers Grimm, which provide the basis for many of the fairy tales that are part of European traditions, the evil character is often signified by means of their physical differences”. So the fairy tales use physical disability to represent characteristic or moral shortcomings. Fairy tales can be used in a more positive manner. Openshaw (2012: 83) makes the link when she proposes the use of fairy tales in social work: “Reading stories, especially fairy tales, can help teach social skills. Whenever the characters behave in an inappropriate manner, the school social worker should stop and ask why the behavior is wrong and what the character should do instead”. Having people write stories, or fairy tales, to tell their own stories is also a useful tool in social work, as important information can be gleaned from such stories, information that the client might not otherwise feel so free to divulge.

I begin with a review of the literature, which illustrates further meanings of the word ‘creativity’. Then I will give a brief explanation of Corpus Linguistics, followed by a description of the methodology used in the study and its outcome; to finish with, pedagogical conclusions are drawn from the project.

2. Creativity

Creative use of language is often cited as one of the aims for our students: to encourage them to practise their creative production of the L2 (Boden 2004; Carter 2004; Pope 2005; Pope and Swann 2011; Tin 2013; Tomlinson 2015; Widodo and Cirocki 2012). Widodo and Cirocki (2012: xx) state that the aim for their book on creativity in EFL is to “foster a scholarly discussion on the issue of innovation and creativity in English language instruction at different levels”. In the descriptors of the Common

European Framework of Reference for Language (Council of Europe 2001), there are descriptors for Creative Writing (*ibid.*: 63). Moreover, in the descriptor for Coherence and Cohesion (*ibid.*: 125) at the C2 level there is the phrase “Can create coherent and cohesive text [...]”. Apart from these instances, the words *create* or *creativity* do not appear anywhere else. This suggests that this discussion has not progressed much.

Batey’s (2012) work calls for a definition of creativity so as to help investigations into creativity and tries to offer a means of measuring this phenomenon. However, he concludes by asking for more work to be done regarding the former and his proposal for the latter is rather cumbersome. Boden (2004: 1) gives a definition of creativity as “the ability to come up with ideas or artefacts that are new, surprising and valuable”, but later (*ibid.*: 2) makes a distinction between “psychological creativity” and “historical creativity”, in which the former is “coming up with a surprising, valuable idea that’s new *to the person* who comes up with it” while the latter is a totally new idea never seen before.

Pope (2005: 41) traces creation and creativity from a religious notion of creating something from nothing (*ex nihilo*) and a sense that once something had been created then it could not be changed to a post-Darwinian concept of “creation as re-creation”. He claims (*ibid.*: 6) that philosophically “during the closing decades of the twentieth century, almost anything to do with ‘creators’, ‘creation’ and ‘creating’ was roundly attacked”. He notes (*ibid.*: 91) that the creative process is usually “prompted, informed or driven by: inspiration, ecstasy, influence or intertextuality”, and that “creativity is understood in terms of game-like constraints and the kind of rule-making and rule-breaking activity that more or less ‘free’ play may entail”. Michalko (2006: xvii) suggests that creativity comes from the inspiration to see something differently: “By changing your perspectives, you expand your possibilities until you see something that you were unable to see before”. Baréz-Brown (2006: 7) plots the way that creativity is obtained: he claims it goes from “insight” to “ideas” to “impact” to “inspiring opportunities”. Catling and Davies (2002: 17) write that “we often restrain ourselves from being as creative as we can. We hold ourselves in a mental straight jacket”.

Constraint is sometimes needed for creativity to occur – indeed constraints could be represented by the rules that Pope talked about above. This ties in with the work of De Bono (2007), in which he sets a series of constraints and the reader is asked to be creative within those boundaries. It is not that the creative process can be expected to occur in a vacuum: in his summary after his comprehensive search through the past and present of creativity, Pope (2005: 191) notes that “Creation is always from ‘something’ and not ‘from nothing’ but always unique and in some sense fresh every time. It is repetition, but with distinct and significant differences”. Relating this to the field that we are interested in, Pope (*ibid.*: 276) states that “language is hailed as a routine yet remarkable creative resource in every linguistic approach”.

Whereas Pope looks at the literary work of geniuses to obtain an idea of what creativity is and Guy Cook (2011: 301) expresses his claim that everyday creativity exists but that “there is also extraordinary creativity, unequally distributed among a very few individuals, to the great benefit of us all”, Carter (2004) suggests that creativity exists almost as a prerequisite in all parts of language and so is a part of everybody’s life. He writes (*ibid.*: 6) that “creativity is a pervasive feature of spoken language exchanges as well as a key component in interpersonal communication, and

that it is a property actively possessed by all speakers and listeners; it is not simply the domain of a few creatively gifted individuals". Following on from this, Jon Cook (2011: 311) states that "the very idea of 'creative language' is an arbitrary construct, not least because it could be argued that all language is creative, and, if this is true, it follows that no significant distinction can be made between language that is and isn't creative". Tomlinson (2015: 108) continues with this idea and brings it closer to the EFL context by stating "Language use is nearly always creative. Therefore, language learning needs to be creative too". Creativity is something which De Bono (2007: 4) firmly believes "is a skill that everyone can learn, practice and use".

When trying to sum up creativity in language, it is possible to use Pope's (2005: 52) idea that "[c]reativity is extra/ordinary, original and fitting, full-filling, in(ter)ventive, co-operative un/conscious, fe><male¹, re...creation" and also add, as he does, that it should be "creativities are" rather than "creativity is" so as to emphasize the plurality of the construct. Regarding recreation, Tin (2013) states that a problem in EFL is that the communication of meaning is often just the communication of a "meaning" that is already known. She argues that this does not encourage the students to use creative language, but that they tend to rely on tried and tested language that they are safe with. Students need to be pushed into using new ideas and she suggests (2013: 386) that "the use of multicultural experiences and constraints can facilitate creative language use". The idea of constraints leads us back to the ideas of Pope cited above of "game-like constraints" and of De Bono. In this investigation, the constraint put on the learners is the use of specific words and phrases (word clusters) that come from a specific language corpus: they have to work within the constraints set to produce language and therefore meaning that is outside their usual field. Tin presumes that by encouraging the use of new ideas this pushes students into the Zone of Proximal Development and so encourages progression in the L2.

This work is based on the notion of linguistic creativity as something that occurs in everyday language all the time, but that an extra level of creativity can be encouraged and learnt. Rather than attacking creativity, as Pope noted has happened in the past, we must welcome it.

3. Corpus Linguistics

One relatively modern way of looking at language which also allows us to see more of the way language works is Corpus Linguistics, and so this should permit us to "expand" our possibilities, as Michalko (2006) above enticed us to do. However, De Bono (2007: 7) gives a warning about thinking that all there is to it is putting texts into a computer and pushing a key:

Because of the excellence of computers, people are starting to believe that all you need to do is collect data and analyze it. This will give you your decisions, your policies and your strategies. It is an extremely dangerous situation, which will bring progress to a halt.

¹ Pope uses this creative form to indicate it is for both males and females.

There is a huge need to interpret data in different ways; to combine data to design value delivery; to form hypothesis and speculations etc.

It is useful now to explore what is meant by Corpus Linguistics. McEnery *et al.* (2006) date the term to the early 1980s, although its history can be traced to as early as 1940. They state that “in modern linguistics, a corpus can be defined as a body of naturally occurring language” but they also cite Sinclair (1996) when he wrote “a corpus is a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language”. McEnery *et al.* (*ibid.*: 3) add to this that “there is an increasing consensus that a corpus is a collection of machine readable, authentic texts”. They also note (*ibid.*: 15) that there are two main types of corpora: *general* and *specialized* corpora.

Corpora are often used to help find frequency, phrase structure, and collocation (Baker 2006; Baker *et al.* 2006; McEnery *et al.* 2006; O’Keeffe *et al.* 2007; Anderson and Corbett 2009). A part of this is that corpus software also allows us to find clusters of words that frequently appear together, what Biber calls “lexical bundles” (Baker *et al.* 2006: 34). Chen and Baker (2010) as well as McEnery *et al.* (2006) point out the problem of terminology here with different authors using different terms for the same or similar meanings; for example, “routine formulae” (Columas 1979); “lexical phrases” (Nattinger and Decarrico 1992); recurrent word-combinations (Altenberg 1998); “chunks (de Cock 2000); “formulaic sequences” (Wray 2002); n-grams (Stubbs 2007). These can relate to Lewis’s (1997: 33) idea that “the importance of semi-fixed expressions cannot be overestimated” and that “several linguists who have studied and clarified expressions have come to the conclusion that they consist of between two and seven words”. Lewis (*ibid.*: 7) continues by stating that “[t]he Lexical Approach argues that language consists of chunks which, when combined, produce continuous coherent text”. These chunks can be found using appropriate software to analyse a corpus. Analysis can also be carried out to identify the keywords within a corpus. One definition of keywords (Baker *et al.* 2006: 98) is “Any word which is considered ‘focal’ in a text, but not through statistical measures”.

To identify keywords within a text it is useful to have a stop-list, which is a list of words that you wish to eliminate from the word list from your corpus. The most frequent words in almost any English corpus will tend to be the same: for example, the top ten most frequent words in the British National Corpus (BNC) are as follows: *the, be, of, and, a, in, to* (infinitive marker), *have, it, to* (preposition). Finding that these are the same most frequent words in the corpus you have created does not tell you much about what is special about the language of that corpus and so it makes sense to eliminate these words and try to concentrate on the specialized words. These words and word clusters obtained from the specific corpus can work as the constraints that were mentioned in section 2 if students are asked to use them in a task. By explaining where these words and word clusters come from, the students realize that they are dealing with authentic language, even though it might be in a form that they are not used to (i.e. not in a piece of continuous text).

The idea that Corpus Linguistics should aid learning and learners is expressed throughout the literature (e.g. Sinclair 2004; O’Keeffe *et al.* 2007). However, McEnery

et al. (2006: 195) write that “the usefulness of corpora in language pedagogy is an area of ongoing debate”. This article joins in on this debate.

4. Description of the task and its methodology

This project was carried out at the University of Calabria, in the south of Italy, and involved volunteer third-year students (about 21 years of age) from the degree course in Social Work who were attending my ESP (English for Specific Purposes) course and whose knowledge of the English language ranged from the A2 to the B2 levels. The project consisted in the construction of specific word and word-cluster lists that the students could then use to help them write stories. The aim was to be creative in storytelling and to produce something new, not “something out of nothing” but “something from something else”. It also functioned to introduce students to corpus linguistics with the hope that this might be beneficial to them in later studies, as its methods and tools can enable them to track the language of their professional genres and interactions. I decided to use the works of the Brothers Grimm, which are well-known and provide a referent against which to contrast the creativity of students. Jacob Ludwig Carl and Wilhelm Carl Grimm were “German brothers famous for their classic collections of folk songs and folktales, especially for *Kinder- und Hausmärchen* (1812-22; generally known as *Grimm’s Fairy Tales*)”². As stated above, fairy tales can play an important role in social work as they are culturally relevant stories that allow people to identify specific behaviour and can lead to greater understanding of culturally determined reasoning.

To compile the corpus of this project I used the Philip Pullman (2012) book of Grimm Tales. Pullman’s intention (*ibid.*: xiii) in retelling these stories was to “tell the best and most interesting of them, clearing out of the way anything that would prevent them from running freely [...] to produce a version that was as clear as water” and do so in a modern form. As a writer, Pullman felt free to make any changes he felt necessary to achieve his aim. There is no single, unique definitive version of the Grimm tales. Pullman explains in his introduction that this is because, as part of Germany’s oral folk tradition, they are always being interpreted by the teller. This corpus (hereafter called the Grimm Corpus) had to be scanned from the book and then read by an Optical Character Recognition program, in this case ReadIris 11, to convert the scanned text into computer readable text, as McEnery *et al.* suggest corpora should be. This made the data analysable. The software program used was *AntConc* (Anthony 2014).

To create the stop-list for this task I used an adapted version of the top 2,000 words in the BNC. The list was adapted as I took out specific years (the BNC was created in the early 1990s and this is shown by the fact that years such as 1990, 1989, etc. appear in the list). I also removed specific proper nouns for people (including my own name). It is interesting to note that this meant removing 19 names for males and only one name for a woman (that name was Mary; Jesus was also taken out). Individual letters were

² Grimm, Jacob Ludwig Carl and Wilhelm Carl (2008). *Encyclopaedia Britannica 2008 Ultimate Reference Suite*. Chicago: Encyclopædia Britannica.

also eliminated, except for the article *a*. The remaining 2,000 words then acted as my stop-list.

A frequency analysis of the complete texts from the Grimm Corpus gives the following outcome for the top ten most frequent words (see Table 1).

position	frequency	word		position	frequency	word
1	8035	the		6	1613	she
2	5270	and		7	1574	you
3	2960	to		8	1559	of
4	2441	he		9	1555	was
5	2371	a		10	1532	it

Table 1. Top ten words from the Grimm Corpus

These are not very different from the top ten words from the British National Corpus, which can be found online³. Such a list therefore does not help us identify the specific words from this corpus. By applying the stop-list there is a very different picture of the top twenty words: see Table 2.

position	frequency	word		position	frequency	Word
1	164	princess		11	73	castle
2	140	Hans		12	72	witch
3	123	tailor		13	68	deer
4	116	golden		14	65	palace
5	92	dear		15	61	flew
6	91	bird		16	58	wild
7	90	till		17	55	bread
8	83	soldier		18	54	devil
9	77	snow		19	54	majesty
10	73	asleep		20	53	miller

Table 2. Top twenty words after stop-list

This list gives a clearer picture of the Grimm tales; these words can therefore be considered keywords in the tales. I did not analyse their distribution in the texts. Williams (1983) in the introduction to his revised edition of his 1976 work on Keywords writes about how the choice of focal Keywords can seem arbitrary. In this investigation, there has been no selecting of a specific type of word to be investigated (some articles look at only nouns, some adverbs, some cut out all functional language). In this

³ Downloaded from <http://www.kilgariff.co.uk/BNClists/lemma.num> <http://www.kilgariff.co.uk/BNClists/lemma.num> (accessed 1/02/2014).

investigation, any word that survived the stop-list has been included. This means that words such as “himself” and “let” can appear in the list of keywords. These words might not be idiosyncratically key to this particular corpus but do reflect the words used in it. This avoids the arbitrariness of preselecting specific word categories.

In a similar study (Robinson 2010) that used the 1884 Margaret Hunt translation of 209 tales collected by the Brothers Grimm, I obtained the following top twelve results reported here in Table 3:

position	word		position	word
1	king		7	answered
2	saw		8	cried
3	once		9	let
4	himself		10	forest
5	father		11	son
6	daughter		12	wife

Table 3. Top twelve words after stop-list in Robinson 2010 study

This indicates the changes that Pullman has made in his retelling of the stories and could be of interest in further studies.

As noted above, language is not only composed of individual words, but mainly of clusters of words. The *AntConc* program also allows these to be identified; they are referred to as *N-Grams*. The program was set to find any three-, four-, five-, six- or seven-word combination that co-occurred frequently in the corpus. This provided abundant data. For ease of study, I decided to analyse only those clusters that appeared at least seven times, otherwise the number of three-word clusters would have become too big to manage easily. Obviously, a long cluster has within it several smaller word clusters, e.g. the cluster “if only I could get the shivers” includes the six-word clusters “if only I could get the” and “only I could get the shivers”, as well as the five-word clusters of “if only I could get”, “only I could get the”, “I could get the shivers” and so on. These sub-clusters were eliminated (some similar clusters appear in the results, for example “I don’t know what” and “I don’t know” as the latter was not a sub-cluster of the first, but a different cluster). The longest repetition found was “mirror mirror on the wall who in this land is the fairest of all”. Using Lewis’s idea that chunks usually do not exceed seven words, this was the limit that was imposed here and so these longer clusters were not considered. The clusters are displayed in Tables 1a-5a (see Appendix). O’Keeffe *et al.* (2007: 65) note that in many systems the chunks are calculated by the computer “counting characters and spaces only” and so “*it’s* and *don’t* are considered as one ‘word’”. One advantage of creating one’s own small corpus is that this disadvantage can be overcome and so, for example, the word chunk *I don’t know* can be counted as a four-word chunk. This was done here so as to avoid complications and discussions with learners who are used to writing for exams in which contraction forms are counted as the constituent parts and not as a single unit. Obviously, not all the clusters have been shown and cut-off points have been decided on frequency: thirty for the three-word

clusters and the top twenty for the four-word clusters. These words and phrases were then typed on a sheet of paper to create a worksheet which was given to students, who were asked to use some of these words and word clusters to help them write a story and to underline these as they progressed, without specifying what type of story to write.

There was a total of twenty-eight stories: some students worked together. These stories totalled 3,483 words in all, 272 of which were single words from the work sheet (including repetitions of the same word). There were 106 three-word clusters, 65 four-word clusters, 15 five-word clusters and two six-word clusters. As for the percentages of all words written, 7.8% of the total were represented by single words from the worksheet, three-word clusters accounted for 9.1% of the whole, four-word clusters amounted to 7.45% and the five-word clusters made up 2.15% of the total. In fact, the worksheet provided 26.7% of the words in the final stories. Students managed to complete the task and were 'successfully creative', but some variation in the creative quality of their stories could be appreciated as shown by examples (1), (2) and (3). Creativity was detected in three major strands: the collocational recreation of traditional discursive formulas, as can be seen in the sketchy beginning of (1): "There was once a princess. She lived in a beautiful castle. The princess...", the setting up of novel scenarios, such as a "big modern palace" (2) or "a picnic in the park" (3), and unexpected storyline twists, present in all three examples.

Here is a first story example: the parts underlined are the words or word clusters that were used from the worksheet.

(1) there was once a Princess. She lived in beautiful castle. The Princess saw a wild animal in the forest. She is rescued by a young man with a little red riding hood.

Although very short and to the point, (1) does use a high percentage of elements from the worksheet, 15 out of 32 words. It is short and to the point. It is creative in that it is a new story and shows a creative twist at the end, where it is the young male rescuer who is wearing a "little red riding hood" and not a young girl as in more traditional stories. The constraints have worked to help forge the story in which we have a clear, romantic setting, a dangerous threat, and a resolution through heroic deeds.

Here is a slightly longer story.

(2) once upon a time there was a Princess in a big modern Palace. She always wanted to eat bread because she felt alone. One day came into this Palace a soldier: he was so beautiful that the Princess fell in love with him immediately. The young man was excited because he couldn't believe the Princess love him. Unfortunately there were the problem: the soldier hate bread because he couldn't eat it.

The king and queen, father and mother of Princess try to search solution: the two lovers can stay together only if the Princess eat bread under the Juniper tree. They accepted this solution because that is a good idea. At the end they married and divorced after two months of marriage.

This creative piece of writing uses elements from the worksheet and from traditional fairy tales in unexpected ways. By being asked to use specific words (our creative constraints) we find a princess who loved eating bread and her lover who could not

stand it. The use of “under the Juniper tree” added another peculiar aspect to this tale (although for some reason juniper trees were quite popular in Pullman’s take on the Grimm stories). Again, we find a creative twist at the end that moves this away from the traditional style of fairy tales.

Here is a final example, also of a brief nature:

(3) today is a beautiful day! The sun is shining. I and my brother Hans go to a picnic in the park near the castle. Hans meets the old woman teacher. I don’t know this lady.

Here almost a quarter of the total words come from the worksheet.

A follow-up questionnaire was used to elicit some responses from the story writers. The questionnaire involved a language level self-evaluation question. In this, twenty students self-evaluated themselves as B1 level, five as being at a B2 level and five at an A2 level, while two self-evaluated as A1.

The writers were asked to grade their answers to four questions on a scale of 1 to 5 (where 1 equals Not at all, and 5 equals A lot). The questions and results are in the following table (Table 4):

Question	Average reply
Did you enjoy writing the story?	4.3
Do you think story writing could be useful for your future job?	2.5
Is creativity important in your future job?	4.6
How creative do you think this activity was for you?	4.2

Table 4. Creativity questions

The first question concerned enjoyment of writing a creative story, and it can be argued that their enjoyment can be equated to a positive learning experience, which was a part of Dörnyei and Ushioda’s (2009) L2 Motivational Self System that was mentioned at the start of this article. The students do not see storytelling as an essential part of their future jobs which, in the case of these students, is to be social workers or social work policy makers, but they do recognize creativity as being important in that future employment. Overall, they saw the activity as being creative, even though they were given constraints to work within (or maybe because of them).

Using the descriptive references in section 1, it can be argued that the goal of creativity has been achieved. Whether or not this then motivates these students to use more Corpus Linguistics and/or motivates them to study more would need a more longitudinal study than this was designed to be.

5. Conclusion and pedagogical implications

This project attempts to show that Corpus Linguistics could easily be used at a small-scale local level, even though the act of creating a corpus might seem daunting for some.

One of the first issues to be addressed is whether creativity is important in EFL, and especially whether or not it is important in ESP. If, as many (see section 1) would say, creativity is an aid to language learning, then it should be encouraged everywhere. This study involved students from a specific ESP class, but they could have been from any. The task was aimed at fairy tales as this genre was used as a (hopefully) fun way to show students what could be achieved by using Corpus Linguistics and their own creativity. It was felt that if students understood that Corpus Linguistics allowed them to produce genre-specific writing in something other than their chosen field of study, then it could be used to advantage in their own ESP training. It would help motivate them as they successfully produce a piece of written work. The link between social services and fairy tales was demonstrated earlier in the introduction. Now these students might be encouraged to use fairy tales, maybe getting social service users to write them (even with the help of this worksheet, or an Italian version of it) as a means of telling their own narratives.

This project was also an exercise in creativity, an idea which has many aspects to it, as this article has attempted to demonstrate. Looking back at Pope's (2005) definition, it is possible to interpret this creative writing as using language that is 'original' in thought or construct and be 'fitting' for that particular genre. It 'full-fills' the need to communicate. The language used was a way to be 'inventive' in creating new stories or new phrases and it allowed people to 'intervene' or interact in a certain communicative discourse. Although, as stated above, creativity can be the creation of something new, what is seen here is the 're-creation' of language already used by others and now used in a new form by a new person. This relates to Boden's (2004) idea of psychological creativity.

To sum up, returning to Baréz-Brown's (2006: 7) idea for creativity (insight, ideas, impact, inspiring opportunities), then it can be considered that 'insight' could connect to the way that a corpus allows us to see the language that is used and how it is used, the 'ideas' can be the fact that from this corpus we can begin to get ideas of what to do with the language, and this has an 'impact' on what we intend to do in a language programme and then it flows on to give 'inspiring opportunities' to the users of the language.

Here, through the use of Corpus Linguistics, the students were given the opportunity to use authentic language to produce a piece of written language in a genre that they were not accustomed to, and this task was successfully accomplished. Dörnyei and Ushioda (2009) suggest that a positive L2 learning experience like this should motivate learners to study more. These learners might be motivated to employ Corpus Linguistics in their own field of study. The experience has attempted to push students into using new ideas, as Tin (2013) suggests, and has tried to make language learning creative as Tomlinson (2015) urged. Finally, this work might have helped answer Widodo and Cirocki's (2012) plea for greater discussion about creativity in EFL and will encourage the debate about the "usefulness of corpora in language pedagogy".

There is of course a lot more to be done. Further research is needed to devise an easily usable clear definition of creativity in language learning and use. This investigation has relied upon personal judgement as to whether the goal of being creative was met or not, but a more scientific approach needs to be introduced.

References

- Altenberg B. 1998. On the phraseology of spoken English: the evidence of recurrent word-combinations. In A. P. Cowie (ed.), *Phraseology: Theory, Analysis and Applications*. Oxford: Oxford University Press. 101-122.
- Anderson W. and J. Corbett 2009. *Exploring English with Online Corpora: An Introduction*. Basingstoke: Palgrave MacMillan.
- Anthony L. 2014. *AntConc* (Version 3.4.3) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved from <http://www.laurenceanthony.net/> (last accessed 18/02/2015).
- Baker P. 2006. *Using Corpora in Discourse Analysis*. London: Continuum.
- Baker P., A. Hardie and T. McEnery 2006. *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Baréz-Brown C. 2006. *How to Have Kick-ass Ideas*. London: Harper Element.
- Batey M. 2012. The measurement of creativity: from definitional consensus to the introduction of a new heuristic framework. *Creativity Research Journal* 24/1: 55-65.
- Boden M.A. 2004. *The Creative Mind: Myths and Mechanisms* (2nd ed.). Abingdon: Routledge.
- Carter R. 2004. *Language and Creativity: The Art of Common Talk*. Abingdon: Routledge.
- Catling T. and M. Davies 2002. *Think!* Chichester: Capstone Books.
- Chen H. and P. Baker 2010. Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology* Volume 14, Number 2, 30-49. <http://llt.msu.edu/vol14num2/chenbaker.pdf> (accessed 11/06/2017).
- Cook G. 2011. In defence of genius. In J. Swann, R. Pope and R. Carter (eds), *Creativity in Language and Literature: The State of the Art*. Basingstoke: Palgrave Macmillan: 290-303.
- Cook J. 2011. Three reflections on creativity and writing. In J. Swann J., R. Pope and R. Carter (eds), *Creativity in Language and Literature: The State of the Art*. Basingstoke: Palgrave Macmillan: 311-322.
- Coulmas F. 1979. On the sociolinguistic relevance of routine formulae. *Journal of Pragmatics*, 3: 239-66.
- Council of Europe 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- de Cock S. 2000. Repetitive phrasal chunkiness and advanced EFL speech and writing. In C. Mair & M. Hundt (eds), *Corpus Linguistics and Linguistic Theory*. Amsterdam Rodopi: 51-68.
- De Bono E. 2007. *How to Have Creative Ideas*. London: Vermillion Press.
- Dörnyei Z. and E. Ushioda (eds) 2009. *Motivation, Language, Identity and the L2 Self*. Bristol: Multilingual Matters.
- Gring-Pemble L. 2003. *Grim Fairy Tales: The Rhetorical Construction of American Welfare Policy*. Westport: Praeger Publishers.
- Lewis M. 1997. *Implementing the Lexical Approach: Putting Theory into Practice*. Boston: Thomson.
- McEnery T., R. Xiao and Y. Tono 2006. *Corpus-based Language Studies: An Advanced Resource Book*. Abingdon: Routledge.
- Michalko M. 2006. *Thinkertoys* (2nded). Berkeley: Ten Speed Press.
- Nattinger J. and J. S. Decarrico 1992. *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

- O'Keeffe A., M. McCarthy and R. Carter 2007. *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press.
- Openshaw L. 2012. *Social Work in Schools: Principles and Practice*. New York: Guilford press.
- Pope R. 2005. *Creativity: Theory, History, Practice*. Abingdon: Routledge.
- Pope R. and J. Swann 2011. Introduction: creativity, language, literature. In J. Swann, R. Pope and R. Carter (eds). *Creativity in Language & Literature: The State of the Art*. Basingstoke: Palgrave Macmillan: 1-22.
- Pullman P. 2012. *Grimm Tales for Young and Old*. London: Penguin Classics.
- Robinson I. 2010. Creativity and Corpus Linguistics: using a corpus to write stories. *Humanising Language Teaching* 2010, Vol. 4, <http://www.hltmag.co.uk/aug10/idea.htm>.
- Sinclair J. 1996. 'EAGLES preliminary recommendations on corpus typology'. EAG-TCWG-CTYP/P. Pisa: ILC-CNR.
- Sinclair J. (ed.). 2004. *How to Use Corpora in Language Teaching*. Amsterdam/Philadelphia: John Benjamins.
- Slessor S. 2014. Working with adults: disability and sensory impairment. In J. Lishman, C. Yuill, J. Brannan and A. Gibson (eds), *Social Work: An Introduction*. London: Sage Publications Ltd: 421-430.
- Stubbs M. 2007. An example of frequent English phraseology: distribution, structures and functions. In R. Facchinetti (ed.), *Corpus Linguistics 25 Years On*. Amsterdam: Rodopi. 89-105.
- Tin T.B. 2013. Towards creativity in ELT: the need to say something new. *ELT Journal* 67/4: 385-397.
- Tomlinson B. 2015. Symposium on creativity for a change: creative use of the coursebook. In T. Pattison (ed.), *IATEFL 2014: Harrogate Conference Selections*. IATEFL: Faversham, UK: 107-108.
- Widodo H.P. and A. Cirocki (eds) 2012. *Innovation and Creativity in ELT Methodology*. New York: Nova Science Publishers.
- Williams R. 1983. *Keywords: A Vocabulary of Culture and Society*. Revised edition. New York: Oxford University Press.
- Wray A. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Appendix

TOP THREE-WORD CLUSTERS FROM THE GRIMM CORPUS

position	frequency	cluster	position	frequency	cluster
1	88	the old woman	17	37	the door and
2	76	I don't	18	36	and when he
3	68	as soon as	19	36	came to a
4	68	out of the	20	36	he said I
5	62	I can't	21	36	into the forest
6	60	there was a	22	36	the king's
7	46	I've got	23	35	don't know
8	46	the young man	24	35	in the forest
9	42	he didn't	25	34	and I'll
10	42	went to the	26	34	she didn't
11	41	back to the	27	33	I'm going
12	41	said the king	28	33	said the boy
13	40	he couldn't	29	30	at once and
14	40	said to the	30	30	in front of
15	39	to the king	31	30	she couldn't
16	38	the king and			

TOP FOUR-WORD CLUSTERS FROM THE GRIMM CORPUS

position	frequency	cluster	position	frequency	cluster
1	32	I'm going to	11	13	she said to the
2	30	I don't know	12	13	there was once a
3	21	and as soon as	13	12	as soon as they
4	19	they came to a	14	12	don't know what
5	18	said the old woman	15	12	he came to the
6	17	little red riding hood	16	12	I've got a
7	16	he went to the	17	12	knocked on the door
8	15	as soon as the	18	12	opened the door and
9	13	he said to the	19	12	the king and queen
10	13	once there was a	20	12	under the juniper tree

TOP FIVE-WORD CLUSTERS FROM THE GRIMM CORPUS

Position	frequency	cluster	position	frequency	cluster
1	11	I don't know what	4	7	in the middle of the
2	10	won't be able to	5	7	that's a good idea
3	9	it wasn't long before	6	7	to the king and said

TOP SIX-WORD CLUSTERS FROM THE GRIMM CORPUS

Position	frequency	cluster	position	frequency	cluster
1	10	no that's not my name	3	7	my sister buried all my bones
2	8	my mother cut my head off	4	7	sister dear your brother's here

TOP SEVEN-WORD CLUSTER FROM THE GRIMM CORPUS

Position	frequency	cluster
1	8	if only I could get the shivers