THE STYLOMETRIC ANALYSIS OF PAPERS PRESENTED AT THE DLCC CONFERENCE IN CIESZYN

Anna Ślósarz

Pedagogical University of Kraków, Podchorążych 2 / 565
aslosarz@up.krakow.pl

Abstract: *The purpose of this paper was to analyze expression styles of 253 conference papers published in the 2009, 2010, 2011, 2012, 2013, 2014, 2015 and 2016 DLCC Conference in Cieszyn proceedings. It was hypothesized that the representatives of similar disciplines expressed themselves in a similar fashion, with the country of origin not being a differentiator. A statistical analysis of the vocabulary of the papers was conducted using the "Stylo" program. The visualizations showed similarities in the expression among 216 authors. The expression styles of the papers were qualitatively analysed. It was found that the DLCC participants have developed common terminology, which is one of the conditions for recognizing Distance Learning Studies as an academic discipline, similar to Media Studies or Knowledge and Communication Studies.*

Keywords: *paper, word, stylometry, space map, cluster.*

PRESUPPOSITIONS

Approaching the stylometric study of the DLCC conference papers, it was hypothesized that language serves not only communication but it also reflects the perception of the world of the group, and it even expresses values and the national spirit of the group (Herder, 1772; Humboldt, 1822). The findings of the American anthropological linguistics representatives have also been referred to, who, according to their understanding of the American Indians languages, believed that every language user utilizes a language which is common in a given society, thus *society speaks through the individual* (Sapir, 1949: 533).

Additionally, two hypotheses of the Tartarus School representatives were applied while analyzing the papers. The first one formulated by Jurij Łotman who argued that the text forms sense and therefore its comprehension is not limited only to a given culture but is of a global significance (1984: 74). The second one formulated by Aleksander Piatigorski who stated that the text message has three

characteristics: it is perpetuated, understandable and is a way of communicating (1975: 115-116).

On the basis of cultural linguistics (Anusiewicz, 1994: 10) as well as the linguistic concept of the world (Bartmiński, 2007), it was assumed that language contains the most important contents of culture. Cognitive linguists (Lakoff & Johnson, 1980) showed that the metaphors reflect the way the world is perceived by the speaker. Therefore, the concepts form the structures and organize the way of thinking and judging, which is passed on to other people (laterally) and onto next generations (longitudinally). Anna Wierzbicka argued that the linguistic analysis of acts of expression is a way of learning about culture (1986) and also that the key words characterize cultures (1997).

In order to identify the most important topics according to conference participants and their shared ways of thinking, the papers were searched for the most commonly used words, then grouped into spatial matrices, hierarchical clusters and the lists of the Most Frequent Words (MFW) was prepared. Finally conclusions were drawn about the way Distance Learning (DL) issues were presented at each conference in the following years respectively and on the whole during the period of time the conferences under study were held.

1. AIM, MATERIAL AND METHODS

The aim of the study was to find out which words and collocations appeared most frequently during the 2009-2016 DLCC conferences. It was hypothesized that the representatives of the same or similar scientific disciplines expressed themselves in a similar fashion on the same problem, with the country of origin not being a differentiator. The research material was a corpus of texts consisting of 253 papers published in successive conference proceedings. The papers were saved as separate text files in UTF-8 (aka Unicode) format. Each paper was tagged with the author(s) names(s) and the year of publication. These tags appeared then on matrices and in dendrograms.

A statistical analysis of the vocabulary and collocations of the papers were performed, using the "Stylo" program and the computational stylistics method (Eder, Rybicki, Kestemont 2014). With the use of the "Stylo" package (Eder et al. 2013), Classic Delta distance and taking into consideration 100 MFW stylometric analyses were conducted. Visualizations of stylistic similarities between input texts were obtained:

  1. Principal Component Analysis (PCA) where covariance plot was presented as a space map[1].

---

[1] Principal component analysis (PCA) invented by Karl Pearson (1901) is an statistical procedure used for transformation a set of observations into a set of values of variables

2. Cluster Analysis (CA) with dendrograms showing hierarchical clustering of the analyzed papers.

T a b l e 1.

Number of papers published in conference proceedings

Year	Liczba referatów
2009	26
2010	22
2011	29
2012	35
2013	46
2014	35
2015	30
2016	30
Total	**253**

S o u r c e : O w n w o r k

Therefore, each paper had its closest "neighbour", which was similar to another paper, and so a map of interrelations was created (E d e r, 2014: 5). The smaller "branches" of the dendrograms linked papers similar in terms of vocabulary used, larger ones – signalled the existence of more significant similarities. Each graph was saved as a graphic file in *jpg* format. Graphs and wordlists were grouped into 9 working folders – one for each year and additional one for the all papers.

Texts were not cleaned out of synsemantic words: articles, particles, prepositions and conjunctions, for example: *a, and, in, of, the, to*. Therefore, an "unpurified" corpus of papers was created; next the lists of MFW were populated; one for each year and one for all the conferences combined. The MFW lists were formed and saved automatically, as well as dendrograms and matrices were, in the corresponding working folders. However, lists of the most frequent words (MFW) calculated by *Stylo* program and presented in this paper do not contain synsemantic parts of language; these were omitted for the sake of clarity. In the centre of the examined conference language, only the most distinctive structures, i.e. the MFW typically used by participants of the DLCC conference were listed. Due to limited space, words and expressions appearing less frequently, though related to the subject matter of the conference, were omitted. These words and expressions assisted with putting newly investigated phenomena in the context of the main ones which had already been described. Additionally, they allowed the DL studies to progress in different directions.

The combination of CA, PCA, and MFW was a way of triangulation of the study results. It allowed for inter-complementing three different ways of presenting data.

---

showed on two, three or more dimensional space. It is used to bringing out clear charts for a dataset what make data easy to exploration and visualization.

2. LIMITATIONS

The study of expression style of the papers prepared by the participants of the DLCC conference is characterized by partial objectivity only because:

1. The distance between texts was measured with the usage of words only (without taking into consideration quotes and inter-textual relations).
2. The method of analysis did not include metaphors, stylistic choices or rare vocabulary while those make the individual expression unique.
3. Many papers were written conjointly: 216 researchers prepared 253 papers, with some researchers participating in writing more than one paper. It was impossible to determine the stylistic contribution of every single author.
4. It was not possible to include videoconferencing, which each year played an important role at the conference, bringing together researchers from different parts of the globe thanks to the use of English language.
5. Only one researcher was an English native speaker, so the expression style of other participants was influenced by academics and researchers, and not by the primary groups like family, neighbours, peers or friends.
6. The analysis did not cover the cultural differences in the expression style on the DL, because exclusively the DLCC conference publications were the research material.

2. THE DLCC CONFERENCES

2.1. Topics

The DLCC conferences have been devoted to the widely understood DL. The speakers referred to courses offered on various platforms, Massive Open Online Courses, Open Educational Resources, Lifelong Learning, informatics competences for teachers and students, e-assessments, as well as e-books, web sites, education in the cloud, robots, social media and educational, technical, economic, ethical and other problems and challenges related to those.
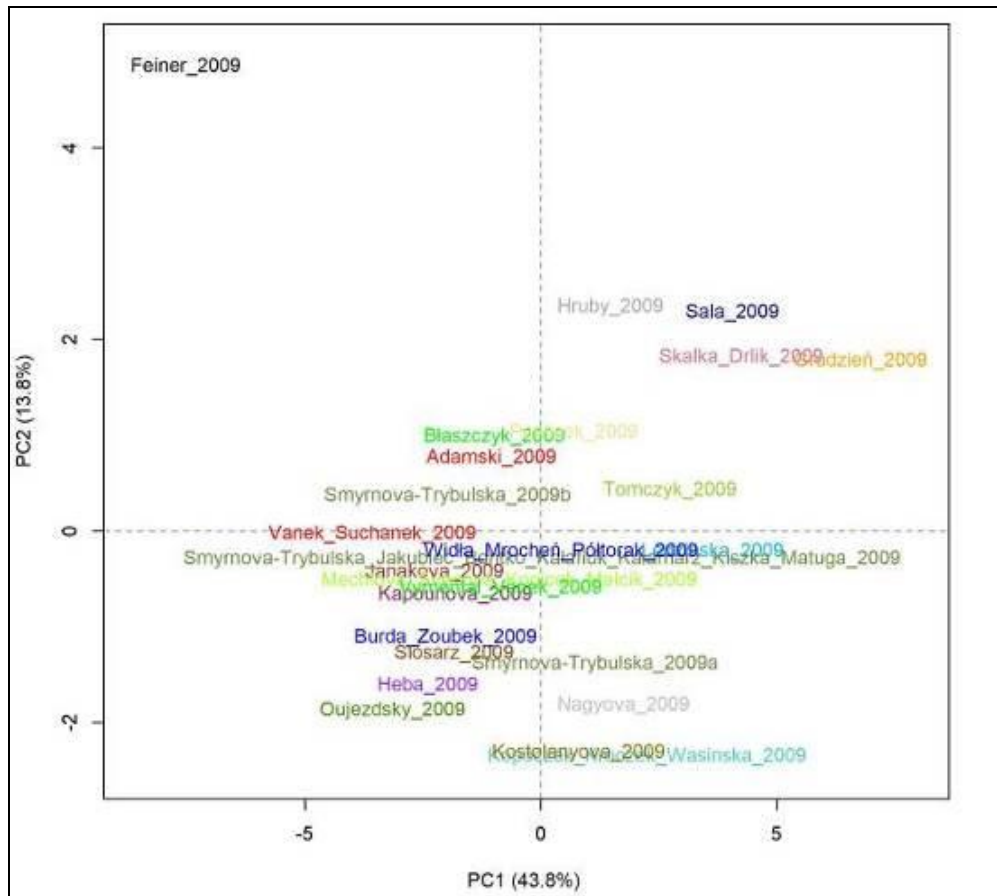
2.2. Participants

For the past eight years the DLCC conference has been gathering professionals from various academic and scientific disciplines, especially those with scientific, humanistic and social background. The participants represented different countries and universities. Among them, there were many computer scientists, including administrators of education platforms. English philologists, teachers, psychologists, librarians, representatives of chemical science, militaries, as well as representatives of linguistics, Roman philology, Polish philology, religious studies, and theology attended the DLCC conferences. It can, therefore, be said that different academic fields and areas were represented by the participants.

3. ANALYSIS OF THE PAPERS

3.1. Year 2009

In 2009, the Principal Component Analysis (PC1) reached about 19pts (taking into consideration 48.9% of data), which represents the highest stylistic variation of the papers in the history of the conference. This seems to be understandable as in 2009 the first DLCC conference was held, and therefore a common style of expression has not yet been developed.
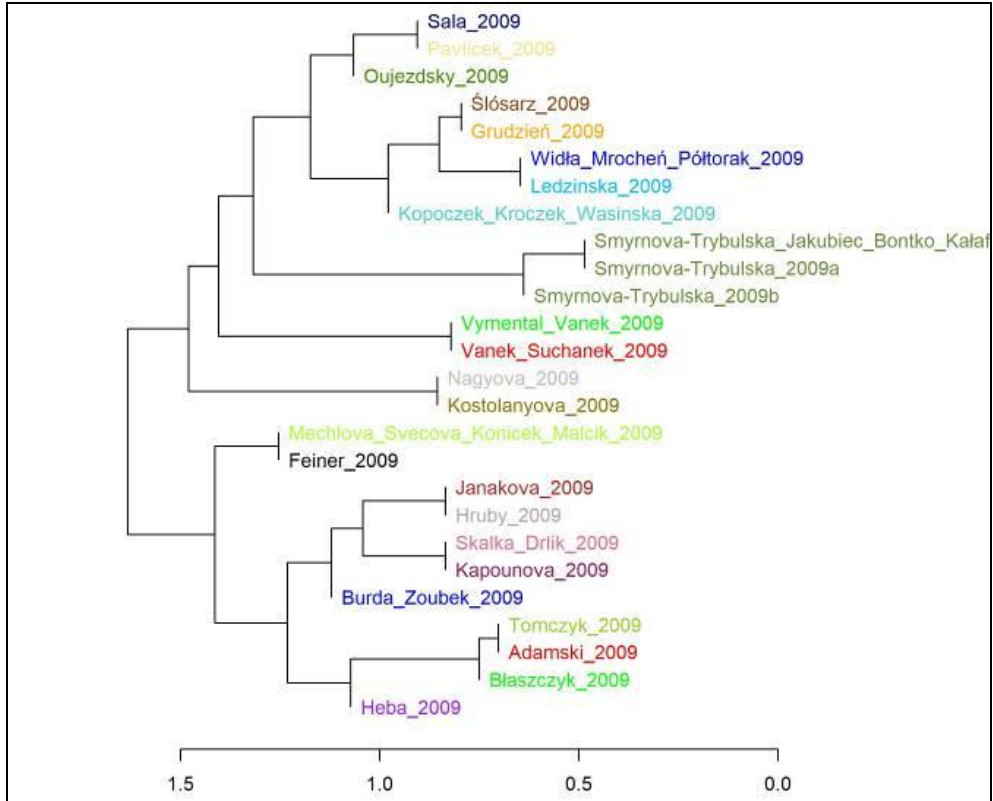

Figure 1. Principal Component Analysis[2] using a covariance matrix (2009)
Source: Own work

---

[2] On all charts Principal Component Analysis refers to correlation of 100 most frequent functional words and shows the position of papers in the space defined by principal components PC1 and PC2 according to K. Pearson and in this case the discriminative strength of words.

On the PCA plot (Figure 1), the names of researchers representing different countries appeared one nest to the other, e.g. Miroslav Hruby (Czech Republik) and Wojciech Sala (Poland) or Ales Oisezdsky (Czech Republic) and Agnieszka Heba (Poland). This means that the country was not a factor differentiating the style of expression.

Figure 2. Cluster Analysis – dendrogram with hierarchical clustering of papers (2009)
Source: Own work

In the upper right corner, there are papers grouped based on the best fitting the stylistics of the conference. Those papers addressed technology: professional teaching at the Higher School of Information Technology (Jacek Grudzień), DL as supporting the education (Wojciech Sala), formal descriptions of electronic questions (Miroslav Hruby) and plagiarism prevention (Jan Skalka and Martin Drlik). The vocabulary of those papers was technology related. At the opposite end, there are papers by Ales Oujezdsky, Katerina Kostolanyova, Agnieszka Heba on multimedia devoted to the development of mathematical competences and by Remigiusz Kopoczek and Katarzyna Kroczek-Wasińska *Artistic Education Based on E-learning*. The positioning of Franz Feiner's paper in the upper left corner of the plot was a result of the fact that it was published in German language.

Cluster Analysis (Figure 2) also revealed some similarities, such as those of E. Smyrnova-Trybulska and dissimilarity of papers, either authored or co-authored by her, with other papers. On the most distant and opposite places, there are once again paper by F. Feiner and by: Erika Machlova et al. (about e-learning at secondary schools), Ingrid Nagyova (*Information iteracy of Pedagogy Students*) and Katerina Kostolanyyova (IT applied in education). Therefore, it can be said that similarities and differences of the papers were illustrated more clearly on the dendrogram thanks to more detailed visualization and more thorough associations than on the PCA plot.

Twenty functional MFW specific for 2009 conference: *students, learning, education, course, distance, information, teaching, can, knowledge, educational, their, use, courses, e-learning, university, multimedia, computer, have, work, materials.*

The 2009 MWFs were mainly associated with the educational process and technological aspects of teaching. They appeared in the plural reflecting the scale of DL usage: *students, courses, materials, their.* The verb *can* reflected the researchers' optimism, similarly positioned statistically next to the verb *can* is the noun *knowledge*, e.g. *students (...) acquiring more knowledge* (Smyrnova-Trybulska, 2009: 9), *students who want to broaden their knowledge* (Widła, Mrochen & Półtorak, 2009: 50). Both words reflected the researchers' belief in the educational possibilities of DL.
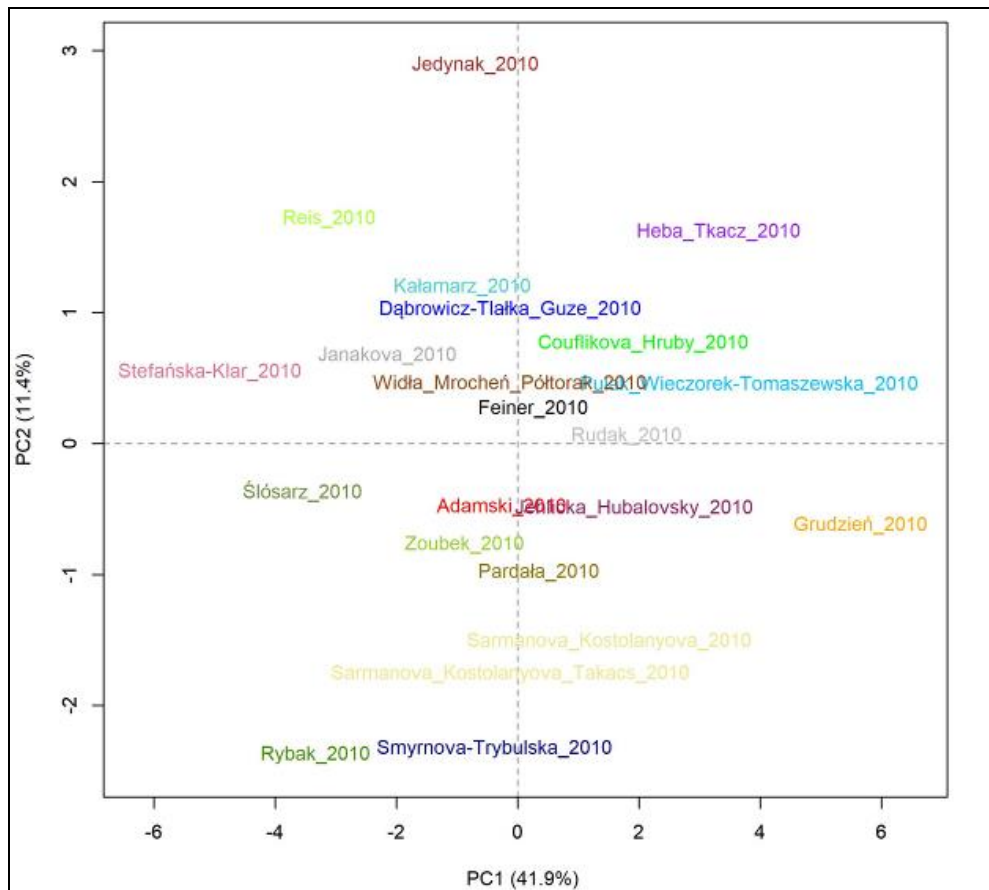
Words *course, courses, learning, teaching, educational* show that the subject matter of the papers focused on the educational process, often personally prepared by the conference participants in the face of technical problems, hence the popularity of nouns such as *computer, materials.*

3.2. Year 2010

The space map for 2010 reveals stylistic differentiation of published papers. They are far from each other, do not form a nucleus as in 2009. Thematic diversity of the texts and addressing the problem from the point of view of different disciplines were also noted. For example, Radosław Jedynak connected supporting of students with remote control software, but Stefańska-Klar connected it with autism and Asperger Syndrome. As a result, these two papers contained different functional vocabulary, so on the plot they were placed relatively far apart.

The PC1 reached about 14pts, i.e. it decreased by 5pts as compared to year 2009. This means that a difficult process of developing a common stylistics of conference language has begun. The 15 authors of the 2009 papers prepared papers for 2010 conference and thus transferred their previous linguistic habits to the 2010 meeting room. To the unification of the expression in 2010 has also contributed publication of the conference proceedings in a timely manner, which allowed the participants to access it at the commencement day of the DLCC conference. Thanks to an excellent organization, the same happened also in 2009. Such a timely publication of conference proceedings was not an easy task for the conference organizers, but

thanks to it the participants were willing to come to the conference counting on the quick acquisition of the points needed for professional stabilization and promotion. At that time, in Poland, English-language publications were scored higher than Polish-language ones.

Figure 3. Principal Component Analysis using a covariance matrix (2010)
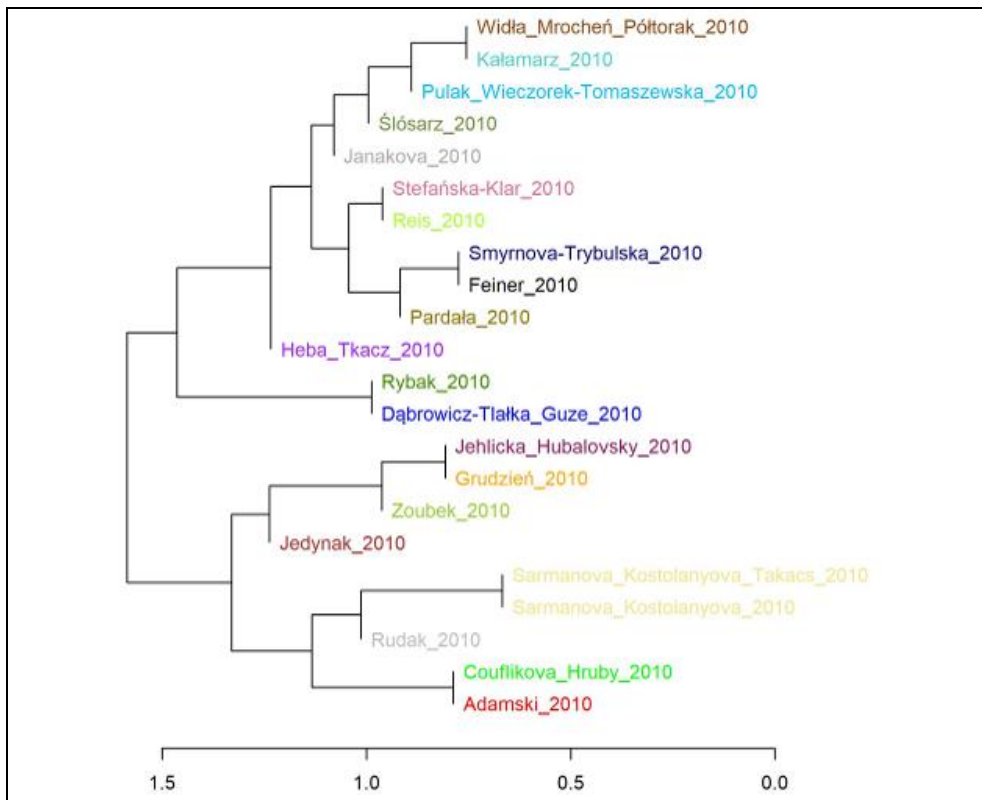*Source: Own work*

Additionally, the papers were published on the Internet, which was a rare occurrence at that time. This was mandated by sponsoring of the first DLCC conference by the Visegrad Fund. That has contributed both to establishing international contacts as well as to the promotion of the DLCC conference and the academic achievements of its participants on the international arena.

The space map (Figure 3) shows that papers by Agnieszka Heba (who participated in the conference for the second time) and Piotr Tkacz were the closest to the general style of the conference. On the other hand, the paper by Anna Rybak, who participated in the debates of the conference only once and prepared her paper independently, was the most distant from style of the DLCC.

Cluster Analysis (Figure 4) revealed stylistic similarities of the papers on various topics. The keynote by Antonio dos Reis, who talked about technical evolution of e-learning, occurred to be stylistically similar to the paper by Renata Stefańska-Klar about *the use of the Internet in supporting and educating persons with autism and Asperger Syndrome* (2010: 249).

Papers by Eugenia Smyrnova-Trybulska and Franz Feiner were also similar. On the other hand, Adam Adamski's original paper *States of consciousness and learning* occurred to be stylistically similar to the paper by Aneta Couflikova and Miroslav Hruby *Protection of learning management systems*. Despite the thematic divergence of both papers, their titles alone reveal stylistic resemblance i.e. the use of the word *learning*.
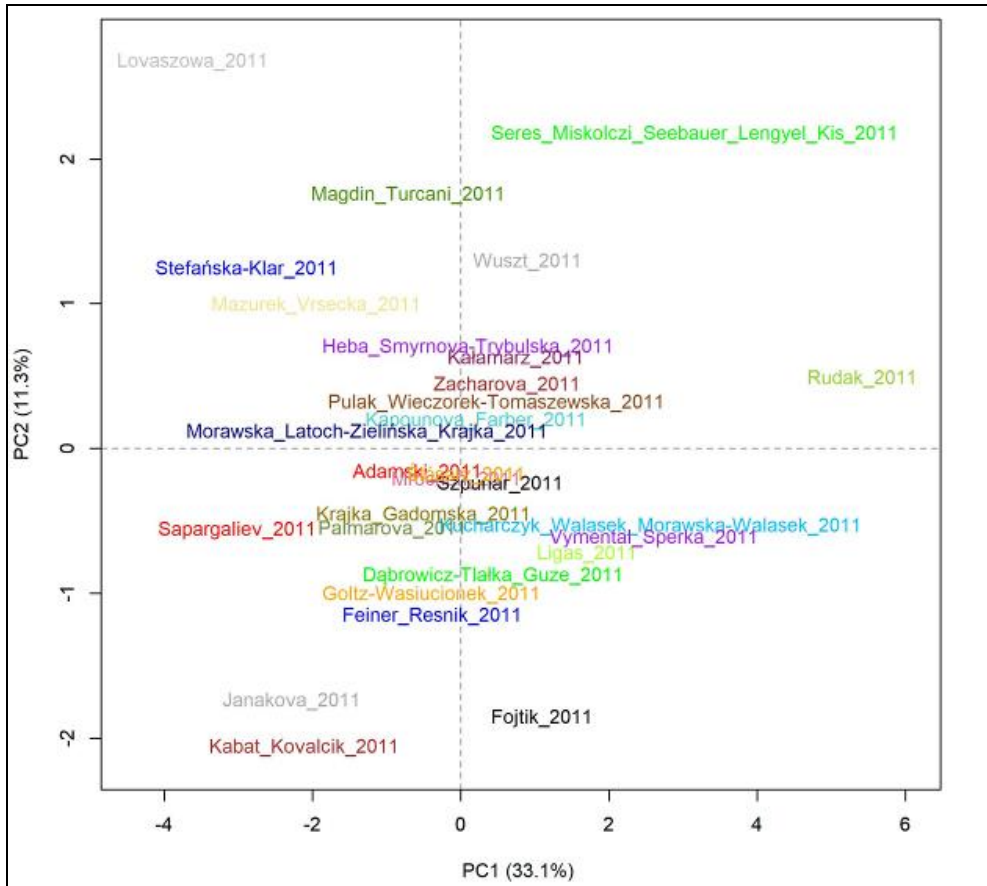

Figure 4. Cluster Analysis – dendrogram with hierarchical clustering of papers (2010)
Source: Own work

Twenty functional MFW specific for 2010 conference: *learning, education, students, teaching, can, course, distance, e-learning, information, their, teacher, educational, new, have, we, internet, student, university, computer, teachers.*

The 2010 MFW, like in the previous year, referred to the educational process. However, their order changed. In the first and the second position appeared the words *learning, education* (in 2009 they were respectively *students, learning*), which meant that the authors limited the debate to their own experiences and made conference discussions more abstract. The word *can* appeared in fifth position (in 2009 it was in eighth position), which, like the word *new*, may suggest more optimistic than a year before researchers' attitude to the capabilities of modern technology
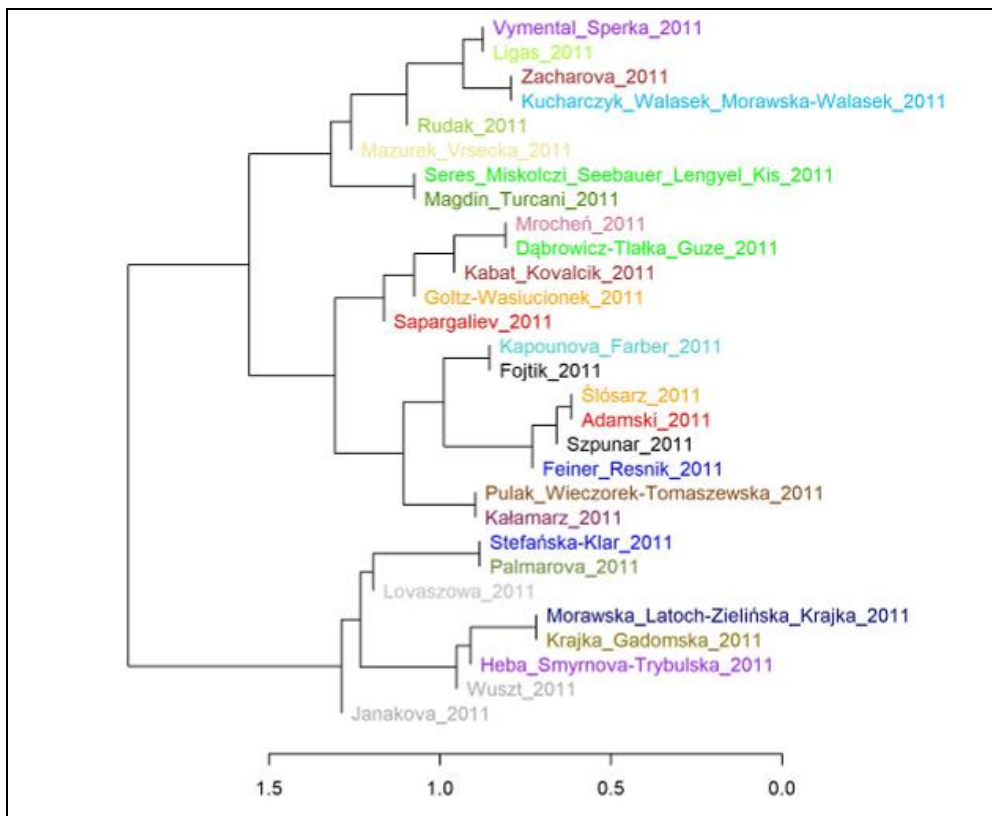
3.3. Year 2011

Figure 5. Principal Component Analysis using a covariance matrix (2011)
Source: Own work

The space map for year 2011 (Figure 5) shows further unifying of the styling of the DLCC conference expression. Many texts are close together. Scale C1 decreased to about 12pts. This can be seen as a manifestation of the linguistic integration of the DL research community, especially that in 2011 out of the 49 authors of the papers as many as 31 (63.3%) participated in the DLCC conference for the first time. Some authors represented countries which so far had not participated in the

conference: Kazakhstan and Hungary. That is why the paper by Danyar Sapargaliyev from Kazakhstan differed from other papers by its individual style of expression. On the other hand, Hungarian researchers (György Seres, Ildiko Miskolczi and others), according to the PC1 co-efficient, adapted the style of their papers to the style of the DLCC conference second best after Leszek Rudak, who participated in the conference for the second time. The cultural conditioning and also diminished accessibility to the conference proceedings the most likely could be the reason for differences in the stylistic expressions between the researchers from Kazakhstan and Hungary and other researchers.

Michal Kabat and Juraj Kovalcik, whose style of expression was the most different from the one conceived at the DLCC conference, participated in the conference for the first time. However, in case of Milena Janakova, the difference, occurring every year, was due to her individual style and the subject matter of her paper.

Figure 6. Cluster Analysis – dendrogram with hierarchical clustering of papers (2011)
Source: Own work

The Cluster Analysis (Figure 6) confirms the distinctiveness of M. Janakova's paper. It also unveils unexpected stylistic similarities to the papers on different topics, such as *Psychological Aspects of Distance Learning in Future Teachers Perceptions. The Role of Prior Experience and Computer Anxiety* by Renata Stefańska-Klar and *Blended Learning of Computer Programming* by Viera Palmarova (significant, that both titles contain words *learning* and *computer*).

Twenty functional MFW specific for 2011 conference: *learning, students, education, can, e-learning, use, their, computer, information, teaching, knowledge, teachers, course, new, have, process, teacher, educational, more, school.*

In 2011, the most popular word was *learning* – same as year before. *Can* appeared in the fourth position instead of the previous year's *teaching,* reflecting an optimistic attitude toward the DL and the tendency to move away from one-way teaching for the sake of establishing rapport – characteristic of connectivism. Replacing *university* by *school* expanded the research area. Other newly occurring words include *use* (6.), *knowledge* (11.), *process* (16), *more* (19.). The word *knowledge* became popular among the participants thanks to the title of the conference and the title of the conference proceedings from 2010: *Use of e-learning in the training of professionals in the knowledge society*. In the top twenty positions did not appear previous year's word *internet* because it has already been an ordinary and more and more easily accessible technology.
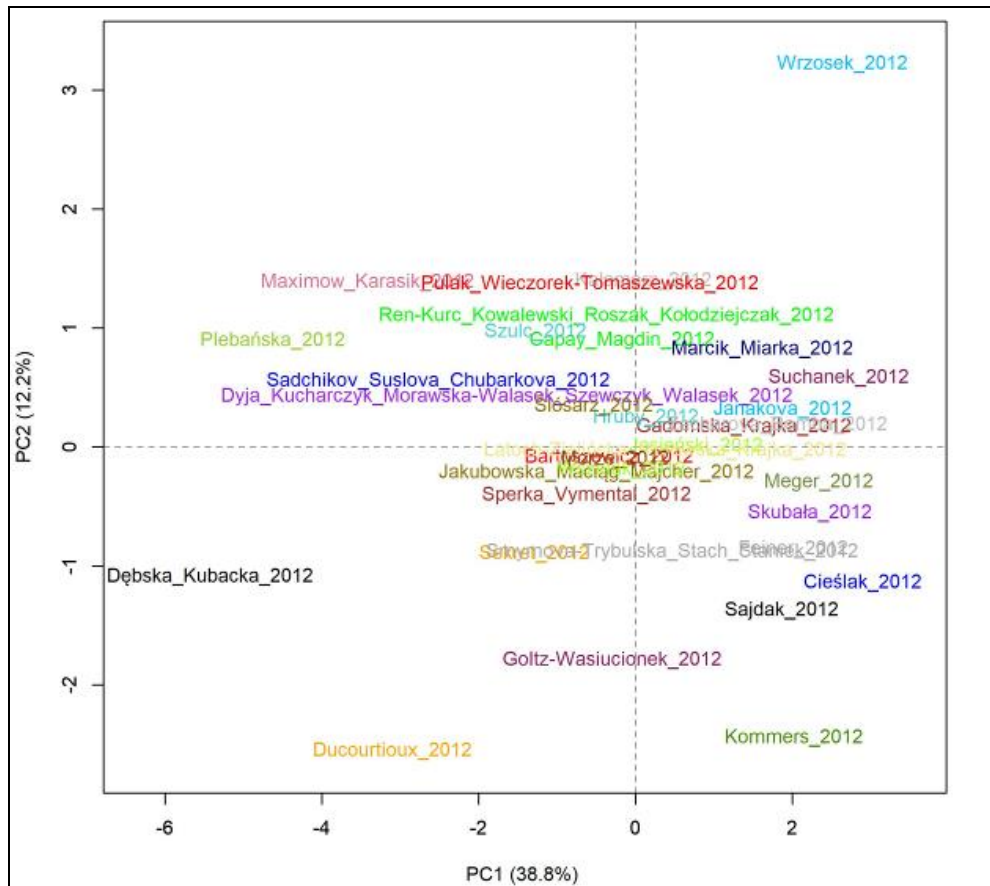
3.4. Year 2012

In 2012, researchers from Russia and Ukraine joined the conference. Among the 57 authors, 36 (63,2%) attended for the first time. Still, the space map shows further unifying of expression style. The PC1 decreased to about 9pts. There were several reasons for this:
1. The researchers attended the DLCC conference for the fourth time.
2. Some researchers were meeting during other parallel conferences (for example: *Virtual University* in Warsaw, *DLSC* in Brno).
3. The conference proceedings were published annually.
4. The papers were made available on the conference website.
5. The Silesian Digital Library (Śląska Biblioteka Cyfrowa) published conference proceedings free of charge.

It is worth noting that on the space map (Figure 7) the texts are mainly placed around positive values, which means that most of them were in line with the style of the conference. Marek Wrzosek, who attended the conference for the first time, did it the best. As a representative of a military academy, he was however familiar with such conferences and with the practical use of the DL.

On the other hand, the least characteristic features of the conference expression style were presented in papers by Sebastien Ducourtioux and by Barbara Dębska & Agnieszka Kubacka, who attended the conference for the first time. In addition,

they came from distant research centres (Dunkirk, France; Krosno, Poland), which had not been previously represented at the conference.

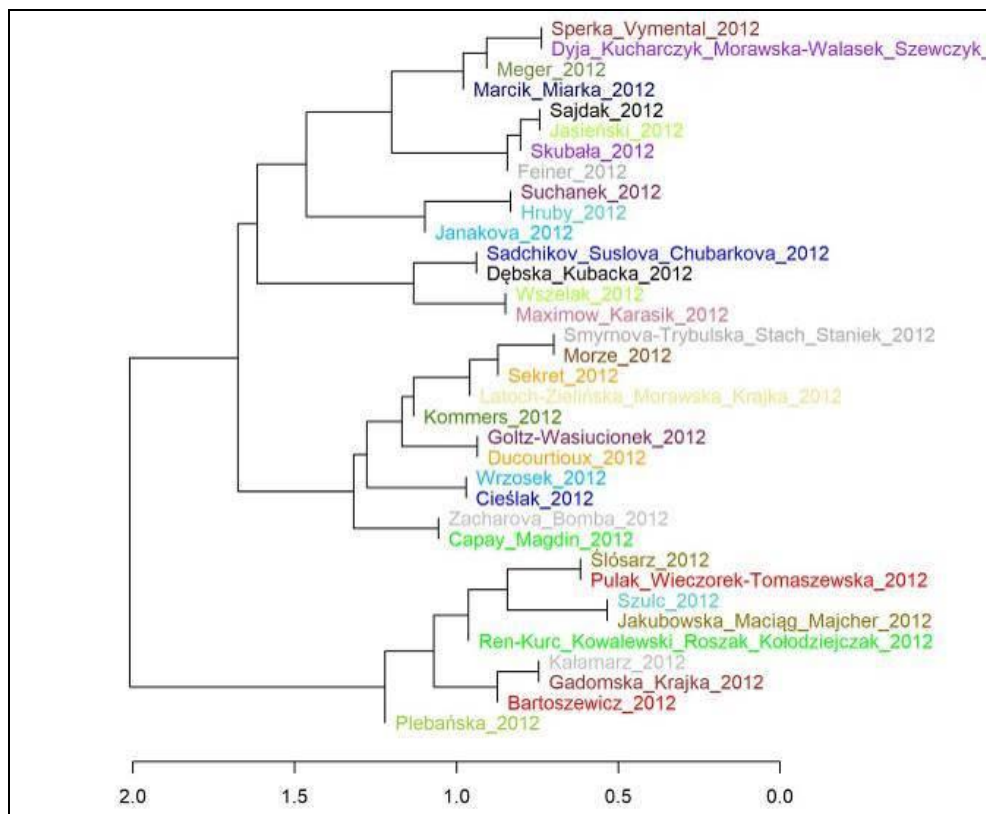**Figure 7. Principal Component Analysis using a covariance matrix (2012)**
Source: Own work

The Cluster Analysis (Figure 8) shows that the seemingly extremely different stylistically papers by S. Ducourtioux and M. Wrzosek, had however a lot in common in terms of deeper meaning.

Twenty functional MFW specific for 2012 conference: *learning, education, students, e-learning, knowledge, information, can, their, educational, process, they, teaching, language, have, was, should, system, development, teacher, competences.*

The MFW for year 2012 found in the first three positions did not change since the beginning of the conference, and they were found in the same positions as in 2010. Newly appearing words were *development* and *competences* – they came up probably in connection with the title of the conference and the conference proceedings from 2011: *Use of e-learning in the development of the key*

competences. On the other hand, the popularity of the word *language* resulted from English philologists (Wojciech Malec, Rüzena Dvorackova) and an Roman philologist (Sebastien Ducourtioux) joining the conference. The new, popular word also turned out to be *system*. The word *their* was replaced by *they,* reflecting moving away from case studies, rapports, didactic and technological issues for the sake of more abstract considerations.

Figure 8. Cluster Analysis – dendrogram with hierarchical clustering of papers (2012)
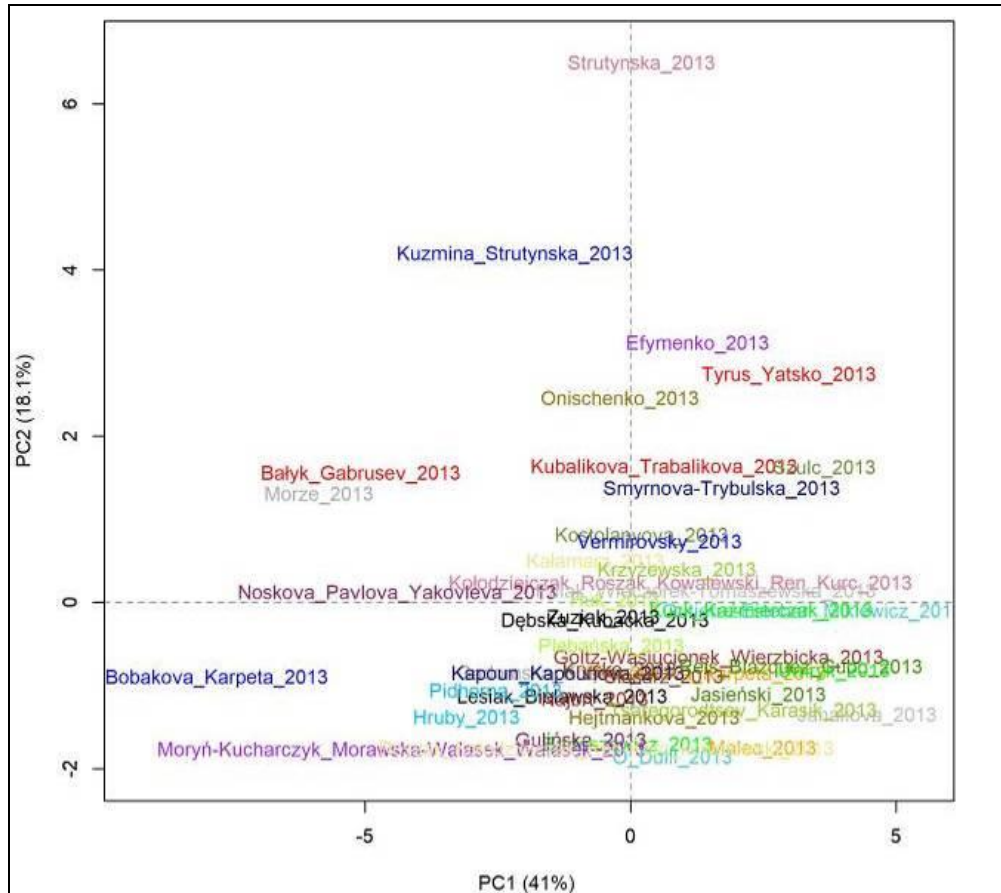
Source: Own work

3.5. 2013

The space map shows that the stylistic expression was varied in 2013. The PC1 reached about 16pts, i.e. 7pts more than in the previous year, but was lower than for the first conference held in 2009, when it reached around 19pts.

Undoubtedly, the factor that differentiated the style of expression was 41 (64%) new participants joining the 2013 conference, from e.g. the Russian Federation, Ukraine, Spain, Great Britain, Bulgaria and Poland. Among them, there were experienced researchers (such as Hanna Gulińska, Tatyana Pavlova), because the

DL had been studied at many research centres. The newcomers brought in their papers their own style of expression to the conference.
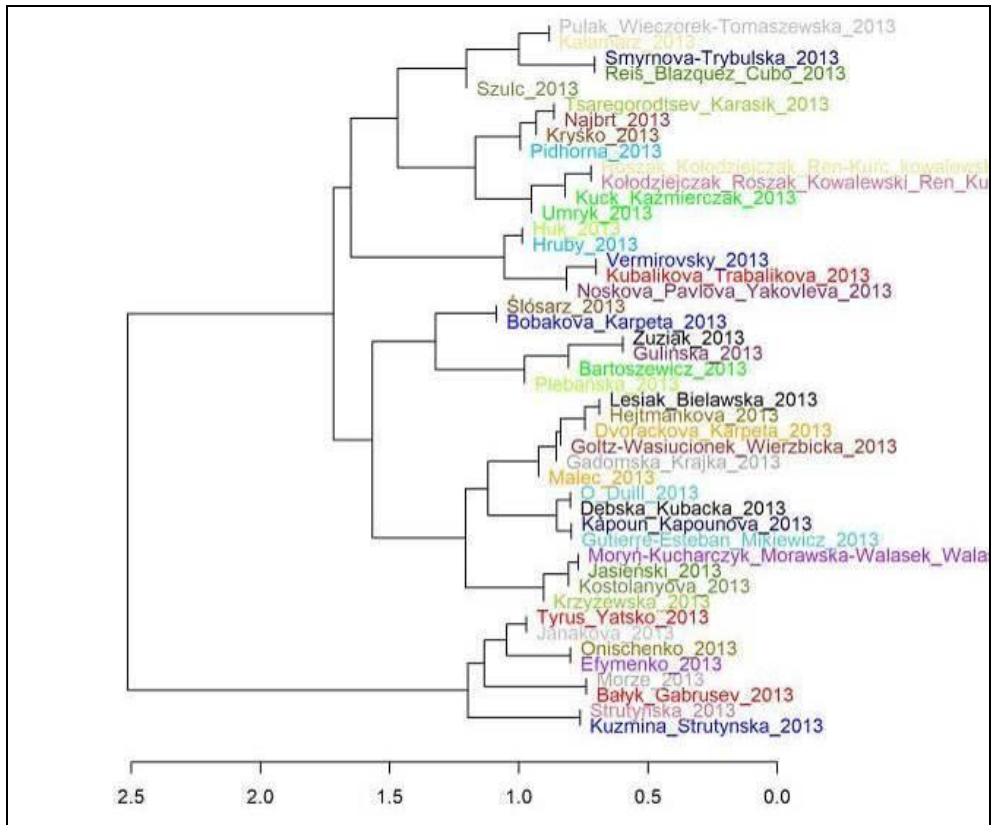

Figure 9. Principal Component Analysis using a covariance matrix (2013)
Source: Own work

The space map shows that many papers were written in the already established style of the DLCC conference. This is also not always true for papers prepared by researchers who joined the DLCC conference in 2013, such as Oksana Strutynska, Natalia Kuzmina, Vasyl Efymenko, Yuri Tyrus, Oksana Yatsko and Sergey Onishchenko from Ukraine.

Dendrogram (Figure 10) reflects stylistic and thematic similarity, e.g. presentations by Hanna Gulińska and Małgorzata Bartoszewicz, who used terminology developed by chemistry scientists. Stylistically similar also occurred to be reports on unrelated topics, e.g. a paper by Hanna Gulińska and by Wojciech Jan Zuziak *Designing and Programing Robots in Contemporary Didactics in Polish Schools*.

Figure 10. Cluster Analysis – dendrogram with hierarchical clustering of papers (2013)

Source: Own work

Twenty functional MFW specific for 2013 conference: *learning, students, information, education, educational, can, e-learning, their, course, teaching, use, process, teachers, distance, knowledge, new, technology, using, they, training.*
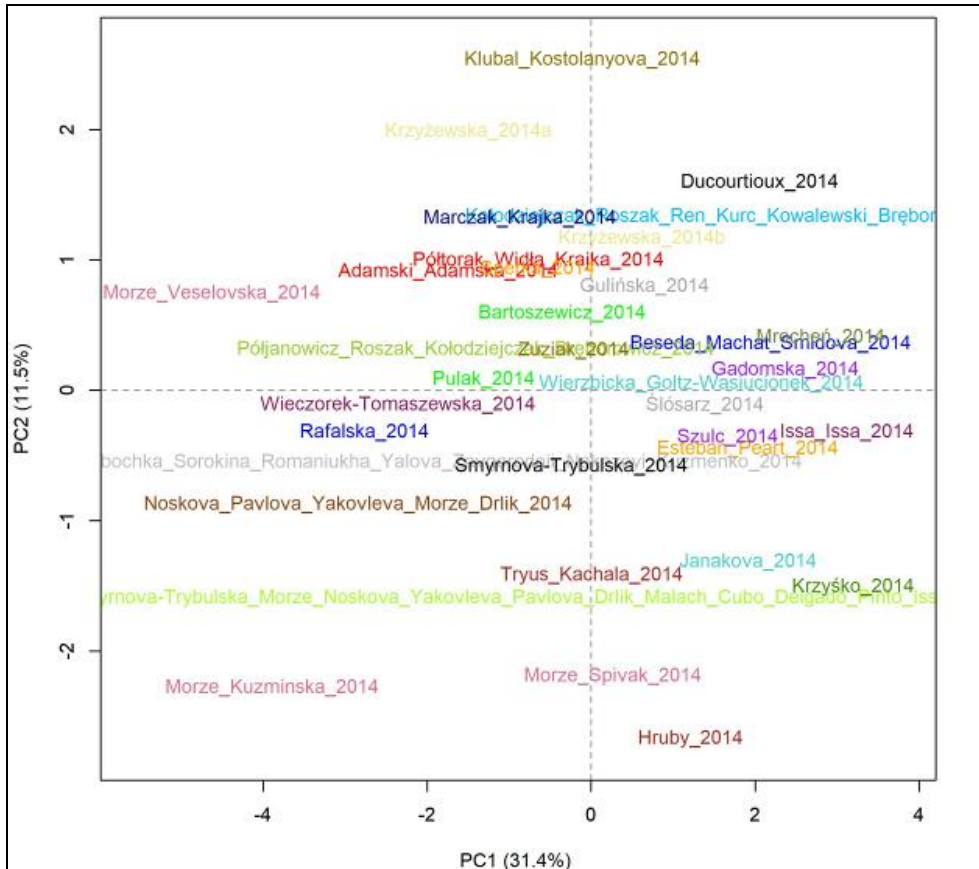
In the vocabulary of the 2013 conference, the importance of information and communication terminology increased: the word *information* emerged in third position, displacing *education*. *Technology* and *training* also emerged, as the latter one was included in the title of the conference and the title of the 2010 conference proceedings: *Use of e-learning in the training of professionals in the knowledge society*. The emergence of such vocabulary was also due to the large number of foreign researchers who joined the conference in 2013 and associated the DL not so much with *learning*, but with ICT and programming.

3.6. Year 2014

The space map for 2014 shows differentiation and further unifying of the expression style of the conference participants. Many of the texts are far from the others, loosely arranged. The PC1 reached about 10pts, which was down by 6pts

when compared to the previous year and almost by half when compared with the first, 2009, conference, when it reached 19pts.
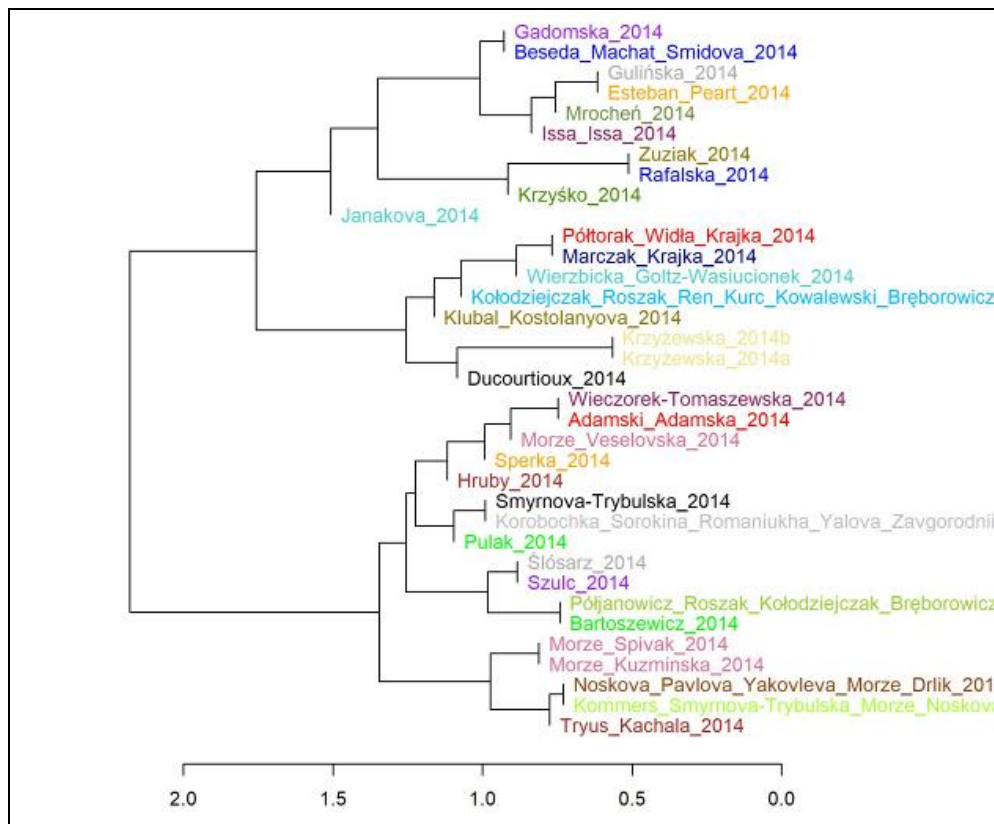

Figure 11. Principal Component Analysis using a covariance matrix (2014)
Source: Own work

The unification of the style of expression at the 2014 DLCC conference was possible for a several factors:

1. In 2014, the conference's tradition was already five-year old and the conference had an impact on 158 participants coming from various research centres; 158 authors and co-authors of papers had attended the conference by 2013, inclusive.
2. Conference materials were published in monographs and made available on the conference website, while scans of monographs were made available by Silesian Digital Library free of charge.
3. The majority of those who joined the conference in 2014 were co-authors of papers being presented.

4. In 2014, the international IRNET[3] project was launched, which allowed for longer lasting academic meetings.

Figure 12 Cluster Analysis – dendrogram with hierarchical clustering of papers (2014)

Source: Own work

In 2014, there were 26 new participants. In the disseminated conference materials, they met with the expression style which had already been well established. Katerina Kostolanyova, who attended the DLCC conference since its very beginning, i.e. since 2009, and Libor Klubal with their papers entered very well in the conference expression style. Sebastian Ducourtioux fitted his paper in the

---

[3] IRNET - *International and Intercultural Research Network for Study and Development of New Tools and Methods for Advanced Pedagogical Science in the Field of ICT Instruments, e-learning Competences.* Duration of the project: 48 months, from 1st January 2014 till 31st December 2017. Project financed by the European Commission under the 7th Framework Programme, within the Marie Curie Actions International Research Staff Exchange Scheme. Grant Agreement No: PIRSES-GA-2013-612536.

conference expression style most accurately, although his paper published two years before was far from the conference expression style.

The papers co-authored by Natalia Morze and Eugenia Smyrnova-Trybulska significantly differed from other papers (exception: the paper by Yurii Tryus and Tamara Kachala) in terms of stylistic expression. Those authors proposed a new stylistic trend initiated by the launching of the IRNET project. That is why on the dendrogram (Figure 12) papers by those authors were placed next to each other, in the lower part of the graph.

Twenty functional MFW specific for 2014 conference include: *learning, students, information, education, their, university, educational, e-learning, use, can, research, course, online, they, knowledge, process, new, work, development, environment.*
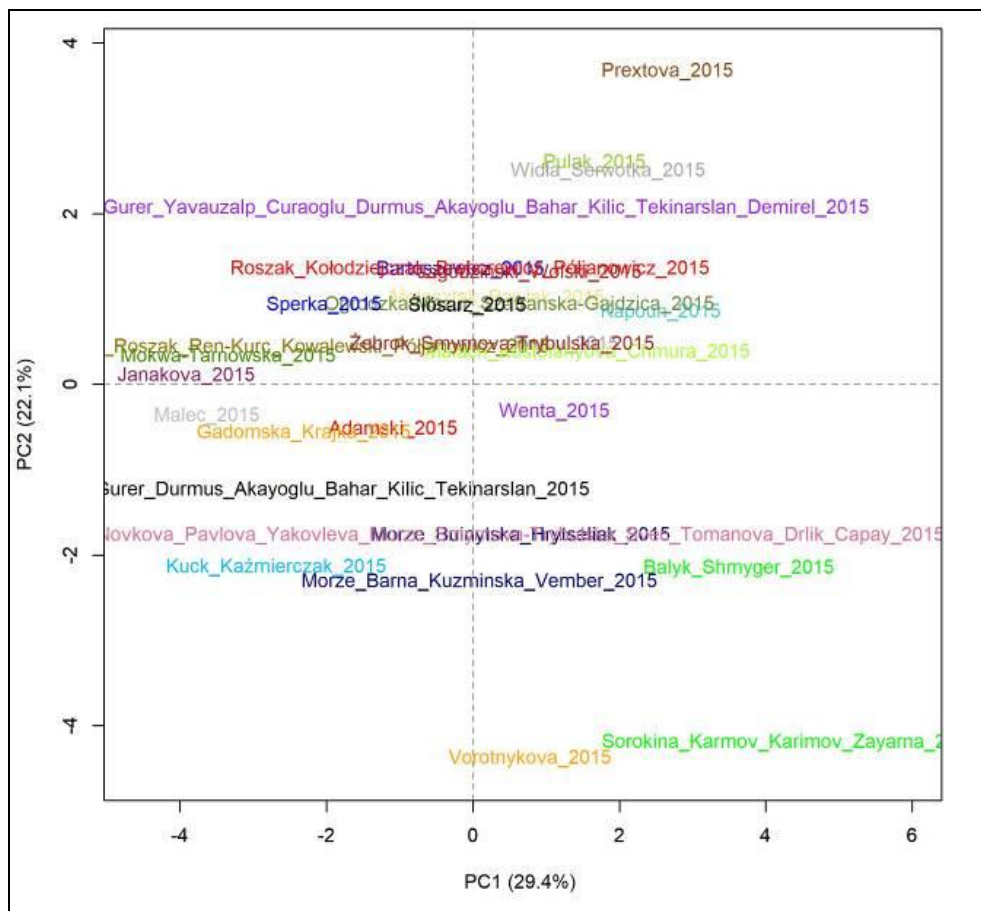
The vocabulary of the 2014 conference papers was similar to the vocabulary of the previous year's conference, which also showed a tendency towards IT. The first four words did not change their positions when compared with the previous year. The word *university* returned and replaced the word *school*. The word *development* came back from 2012, when it was present in both the title of the conference and the title of the conference proceedings: *E-learning and intercultural competencies development in different countries*. The words *research* and *environment* were new to appear, illustrating the conference participants' desire to develop a synthetic academic reflection within the IRNET project, rather than just reporting their own experiences and accomplishments or preparing case studies.

3.7. Year 2015

On the space map for 2015 the texts are arranged in distant groups. PC1 reached about 11pts, which is 1pt more than in the previous year. The higher value of PC2 (9 vs. previous year's 6) showed the diversity of texts written by researchers representing an increasing number of countries and cultures. In 2015, researchers from Turkey joined the conference and with their first co-authored paper they entered well into the expression style and topics of the DLCC conference, discussing their own experience with implementing the DL at their university. Their second text was treating about National ICT Curriculum in Turkey and in this paper they applied an expression style different to the stylistics of the conference. Tatiana Prextova (Czech Republic) similarly to the Turkish researchers participated in the conference for the first time and prepared the paper independently, was the one who did the best with adopting the style of the conference.

Six papers poorly adapted to the conference's style. The most unconforming was the text prepared by Jerzy Kuck and Danuta Kazmierczak, because it addressed business management issues – a topic hardly ever addressed during academic conferences. Two years earlier, the same authors covered the issue of *E-learning as Distance Transfer of Skills and Knowledge*, therefore their paper adopted the

DLCC conference style relatively well (Figure 9, lower right corner of the space map).

Figure 13. Principal Component Analysis using a covariance matrix (2015)
Source: Own work

The Cluster Analysis (Figure 14) showed that the paper discussing the Turkish curriculum, appearing to be written in a unique style and addressing a unique topic, occurred to be similar to the paper by Agnieszka Gadomska and Jarosław Krajka entitled *Pedagogical Solutions in the E-teacher Training – towards Effective Technology Adoption by Teachers*.

Twenty functional MFW specific for 2015 conference: *learning, students, education, information, educational, their, university, can, use, teachers, process, online, course, they, communication, teaching, work, technology, environment, research.*

Figure 14. Cluster Analysis – dendrogram with hierarchical clustering of papers (2015)

Source: Own work

As compared with the year 2014, the new popular words in 2015 were: *teachers, teaching, technology, communication*. The first three words, however, appeared among the MFW in 2013. The only new word in the entire conference cycle turned out to be *communication*. This word refers to a wide spectrum of phenomena, so it turned out to be useful in a situation where researchers represented different countries and cultures, searching for common, therefore very general, terms.

3.8. Year 2016

In 2016, the PC1 reached around 17pts (32.9%), indicating stylistic differences in expression but remained within the limits of the first conference when PC1 reached 19pts (48.9%). This means that despite its global coverage, the expression style of the DLCC conference's participants has become significantly stabilized. Expression styles were distributed relatively evenly across the space map, which indicates their great diversity. Polish researchers were the ones who most precisely adjusted the style of their papers to the style of the conference. Nonetheless, the conference organizer Eugenia Smyrnova-Trybulska and its co-organizer Natalia

Morze initiated new ways of presenting issues by preparing papers that occurred to be most distant from both indicators of stylistic similarity.
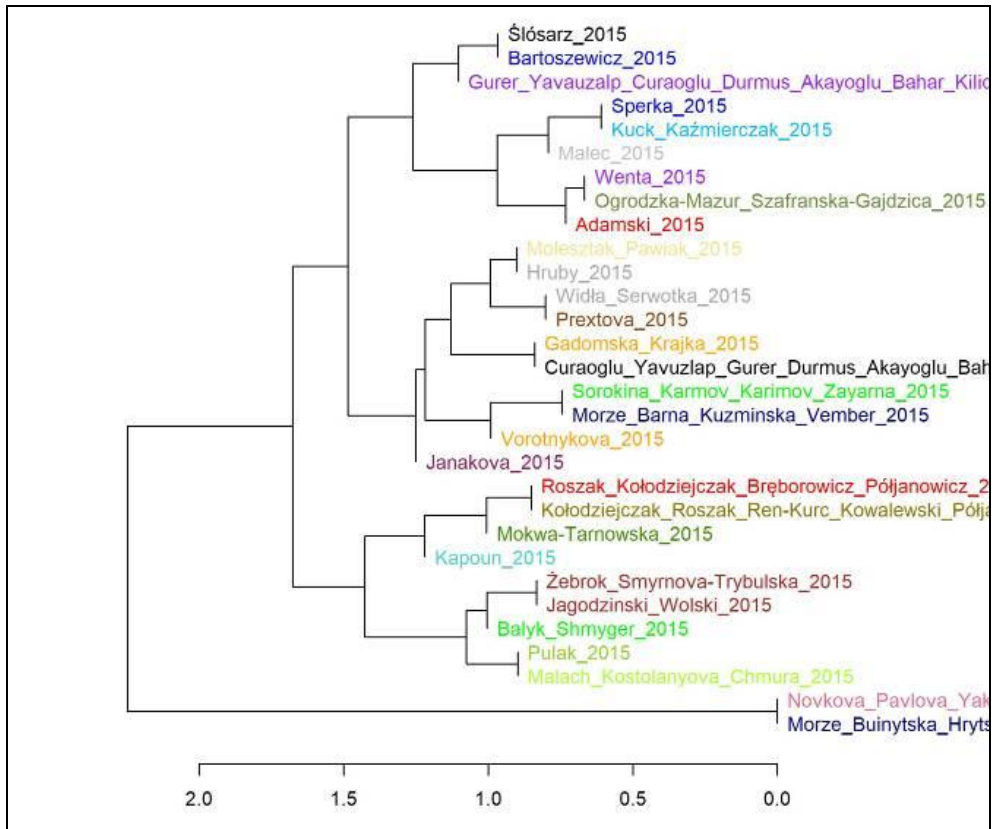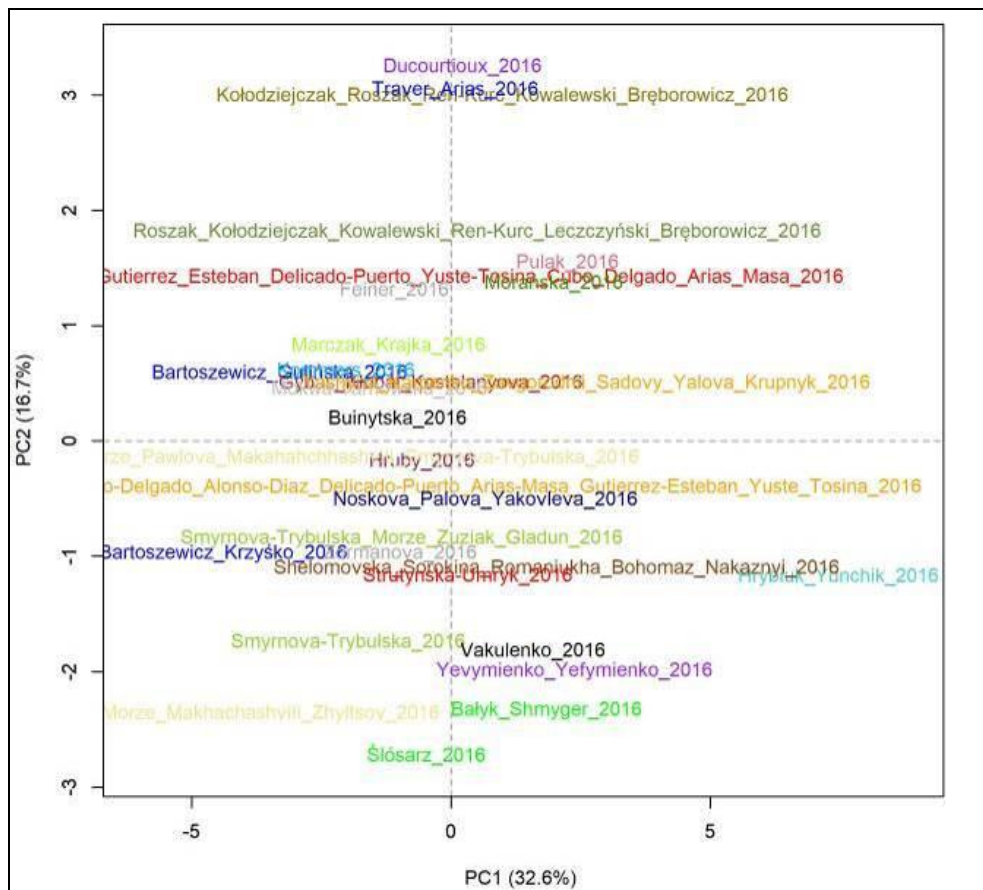

Figure 15. Principal Component Analysis using a covariance matrix (2016)
Source: Own work

The Cluster Analysis (Figure 16) shows that papers by E. Smyrnova-Trybulska and Natalia Morze (a co-author) are placed in the most distant positions, in the area of the upper cluster, which reflects their mutual stylistic similarity, and at the same time confirms distinctness of the topics and expression style when compared to other papers.

Twenty functional MFW specific for 2016 conference: *learning, education, students, research, educational, university, their, work, use, tools, online, information, process, distance, can, teaching, knowledge, teachers, course, figure.*

In the 2016, the order of the MFW was due to the topic of the conference (*E-learning Methodology - Implementation and Evaluation*) and signified increased interest of participants in the research process. The word *research*, which in 2014

appeared for the first time among the most popular words (11th position), in 2016 was found in 4th position. The new words were: *tools, figure*. The former appeared in the title of the conference as well as the title of the 2015 conference proceedings (*IT Tools - Good Practice of Effective Use in Education*). The latter was used for captioning figures and it appeared 254 times in captions; additionally, 13 times it was found in the body of the texts (statistically 8.9 times per one paper). It was observed that the authors illustrated their findings more frequently than in previous years: 2015 – 6.5 figure per paper, 2014 – 4.8, 2013 – 6, 2012 – 7.1, 2011 – 6.3, 2010 – 7.6, 2009 – 3.4. The greater number of visualizations in 2016 resulted from more abstract nature of the reflection, as compared with the previous years, since they referred to the methodology.

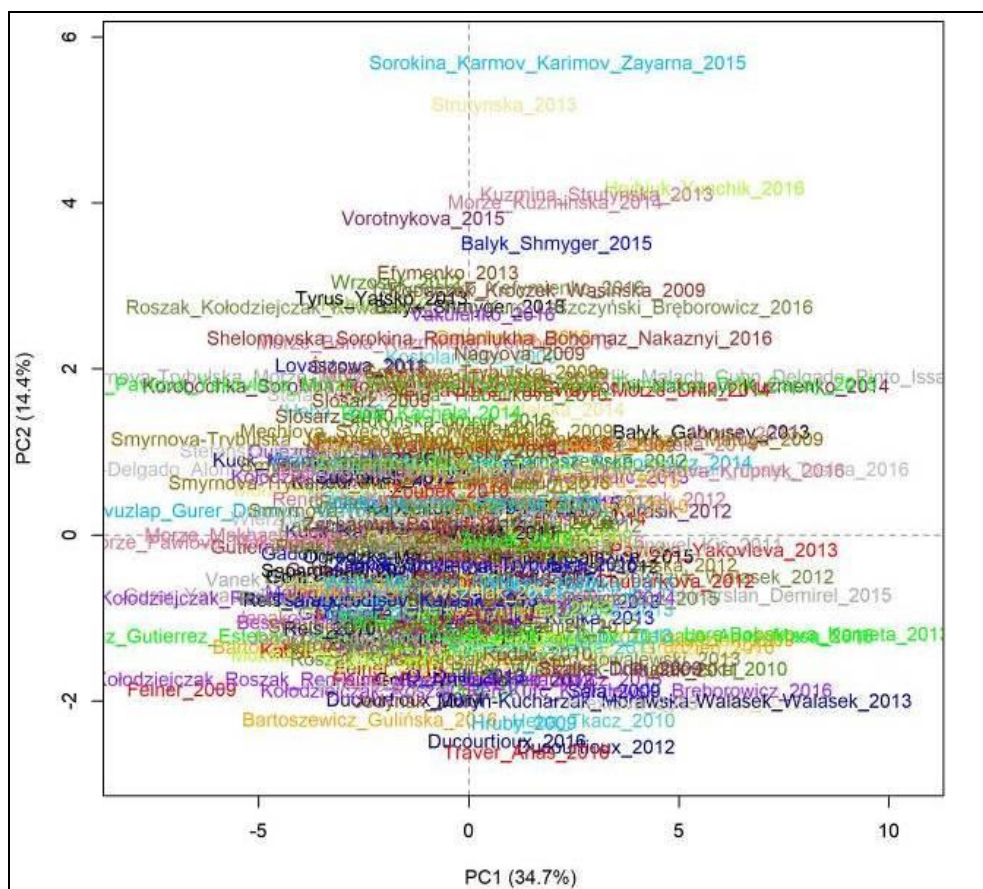**Figure 16. Cluster Analysis – dendrogram with hierarchical clustering of papers (2016)**
Source: Own work

3.9. Years 2009-2016

The space map for the 2009-2016 conferences combined together revealed the highest value of the PC1 when probing for stylistic variations. Its value was about 24pts and 35.6%. The value was due to the thematic diversity of phenomena being

discussed as well as to the geographical span of the conference. Over time, the conference was joined by researchers from a number of countries. Simultaneously, representatives of Europe, Asia, South America, Canada and Australia participated in videoconferences.

Most papers were written in line with the conference style. They are concentrated around the center of the chart. The most distant stylistically was the paper by F. Feiner published in German language. E. Smyrnowa-Trybulska's papers often did not meet the PC1 criterion, proposing new ways of expression. Likewise, the papers whose author or co-author was Hanna Bartoszewicz were characterized by the stylistic distinctiveness resulting from chemical topics they contained.

Figure 17. Principal Component Analysis using a covariance matrix (2009-2016). Due to the large number of participants, legible in this size chart are only the names of authors who have written in the most distinctive style.
Source: Own work

The Cluster Analysis for 2009-2016 conferences is based on a scale 0-7, which is several times greater than the scale for just one conference. It means that although

the style of expression was unified during particular years (thanks to published texts and direct contacts between researchers i.e. by presenting their own papers, by participating in discussions and in meetings during the conference), nonetheless, the papers from the individual years showed differences due to the changing mix of the authors and variety of topics covered in particular years.

The presentation given by Milena Janakova, who discussed, rarely addressed on the DLCC conference, business issues and used a unique expression style, even clearly seen in the title of her 2011 paper *E- learning Applies Attractive Teacher Activities for the Development of Key Competencies in the Field of Information Technology* were found in the most distant, low places of dendrogram.

Figure 18. Cluster Analysis – dendrogram with hierarchical clustering of papers (2009-2016). Due to the large number of participants, in this size chart the names are unreadable. They can be read at a higher resolution.
Source: Own work

Twenty functional MFW specific for 2009-2016 conferences: *learning, students,education, information, their, can, educational, use, teaching, e-learning, course, knowledge, university, process, distance, teachers, they, new, work, research.*

The above listed words refer to the theme of the conference, characterizing it as being dedicated to the problem of DE. Its academic character is reinforced by the words *university, research.*

CONCLUSIONS AND RECOMMENDATIONS

Hypothesis *the representatives of the same disciplines wrote in a similar fashion, with the country of origin not being a differentiator* proved be true in relation to the researched material. Language of academics representing the same, similar and various disciplines, and coming from different cultural backgrounds, also occurred to be alike.

It can be stated that the participants of the DLCC conference, while discussing DL, use a common language regardless of their academic discipline, problem they address, their age, gender, country of origin or cultural background. This common language – being simultaneously developed during other conferences held in Poland, across Europe and worldwide – enables the transfer of experiences and research results, and integrates the DL research community.

Table 2 lists values of the co-efficient of similarity of the papers published between 2009 and 2016.

Table 2.
Values of the co-efficient of similarity of the papers published between 2009 and 2016

Year	PC1	PC2	PC1+PC2
2009	19	7	26
2010	14	6	20
2011	12	5.5	17.5
2012	9	6,5	15.5
2013	16	9	15
2014	10	6	16
2015	11	8.5	19.5
2016	17	6,5	23.5

Source: Own work

Table 2 shows that expressions style of participants of the DLCC conference were becoming more and more stylistically similar from year to year, although the trend fluctuated over the years. A common style was developed despite the increasing geographical span of the conference. New participants were often co-authors, collaborating with their superiors on the preparation of the papers. This means that the DLCCconference affects the development of an international language, allowing to describe the phenomena associated with DL, among the researchers coming from a number of countries.

Therefore, one can say after Edward Sapire that through the individual DL researcher the whole international and multidisciplinary community speaks, because they all use common vocabulary, terminology and expression style. The common language assists with perceiving and discussing the DL-related complex problems and with broadening horizons of the researcher because, according to Ludwig Wittgenstein, *the limits of my language mean the limits of my world* (Wittgenstein 2010: 74).

Developing a common terminology, with an international consensus, is one of the bases for distinguishing an academic discipline [*Rozporządzenie…* 2011 (*Regulation...* 2011)] which has a clearly identified object of research, research methods and social significance. The vocabulary used by researchers places the potential DL academic discipline in the area of social sciences (the Media Studies, Cognitive and Social Communication, Pedagogy, Psychology) as well as technical and mathematical disciplines (Information Technology). The problems addressed during the DLCC conference also show some connection with humanities (Ethnology, Cultural Studies, and Management Studies).

Nonetheless, the DL Studies has not yet been recognized as an academic discipline. The DL research does not secure academic progression; therefore the DL research is conducted gratuitously by passionate academics. Recognizing the DL Studies as an academic discipline would contribute to the development of the DL research, potential of which is expanding and is effectively used to equalizing educational opportunities, individualizing education and economic development in many countries.

REFERENCES

Anusiewicz, J., 1994: *Lingwistyka kulturowa.* Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław 1994, 180 p., ISBN 8322911203.

Bartmiński, J., 2007: *Językowe podstawy obrazu świata.* Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, Lublin, p. 328. ISBN 978-832-272-6853.

Eder, M., 2014: *Metody ścisłe w literaturoznawstwie i pułapki pozornego obiektywizmu – przykład stylometrii.* Teksty Drugie nr 2, 2014, p. 90-105, ISSN 0867-0633.

Eder, M., Kestemont, M. and Rybicki, J. 2013: *Stylometry with R: a suite of tools.* Digital Humanities 2013: Conference Abstracts. Lincoln: University of Nebraska-Lincoln, pp. 487-489.

Eder, M., Rybicki, J. and Kestemont, M. 2014: *'Stylo': a package for stylometric analyses*, https://sites.google.com/site/computationalstylistics/ stylo/stylo_howto.pdf?attredirects=0&d=1, (accessed 4 August 2017).

Herder, J.G., 1770: *Abhandlung über den Ursprung der Sprache*. Christian Friedrich Voß, Berlin 1772, 232 p.

Humboldt, W. von, 1822: *Über das Entstehen der grammatischen Formen und ihren Einfluß auf die Ideenentwicklung*. Abhandlungen der historisch-philologischen Klasse der Königlichen Akademie der Wissenschaften zu Berlin, 1822/1823, Berlin 1825, s. 401-430.

Klar-Stefańska, R., 2010: *Use the Internet in Supporting and Educating Persons with Autism and Asperger Syndrome*. The Use of e-Learning in the Trainingof Proffessionals in the Knowledge Society, ed. E. Smyrnova-Trybulska. University of Silesia in Katowice, Cieszyn 2010, p. 249-260. 346 p. ISBN 978-83-60071-30-4.

Lakoff, G., Johnson, M.: *Metaphors We Live By*. University of Chicago Press, Chicago 1980, 256 p. ISBN 9780226468013.

Pearson, K., (1901): *On Lines and Planes of Closest Fit to Systems of Points in Space*. Philosophical Magazine nr 2 (11), 1901, p. 559-572, http://stat.smmu.edu.cn/history/pearson1901.pdf (accessed 21 August 2017).

Piatigorski, A., 1975: *O możliwościach analizy tekstu jako pewnego typu sygnału*. Semiotyka kultury, ed. E. Janus, M.R. Mayenowa. Państwowy Instytut Wydawniczy, Warszawa 1975, 499 p.

Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 8 sierpnia 2011 r. w sprawie obszarów wiedzy, dziedzin nauki i sztuki oraz dyscyplin naukowych i artystycznych, (2011). Dziennik Ustaw, 30 August, nr 179, pos. 1065, p. 10427-10429, http://www.nauka.gov.pl/g2/oryginal/2013_05/748f846e0255320d2846109ed372a25d.pdf, (accessed 16 Juli 2017).

Sapir, E., 1949: *Speech as a Personality Trait*. Selected Writings of Edward Sapir in Language, Culture and Personality, ed. D.G. Mandelbaum. University of California Press, Berkeley 1963, p. 533-543, 617 p.

Smyrnova-Trybulska, E., 2009. *Introducton*. Theoretical and Practical Aspects of Distance Learning, ed. E. Smyrnova-Trybulska. University of Silesia in Katowice, Cieszyn 2013, p. 9-11. 307 p. ISBN 978-83-925281-4-2.

Widła, H., Mrocheń, I., Półtorak, E., 2009: *Gradual Development of Educatin with the Use of the Methods and Techniques of Distance Learning*. Theoretical and Practical Aspects of Distance Learning, ed. E. Smyrnova-Trybulska. University of Silesia in Katowice, Cieszyn 2013, p. 45-59. 307 p. ISBN 978-83-925281-4-2.

Wierzbicka, A., 1986. *Analiza lingwistyczna aktów mowy jako potencjalny klucz do kultury*. Problemy wiedzy o kulturze, ed. A. Brodzka, M. Hopfinger, J. Lalewicz. Polska Akademia Nauk. Instytut Badań Literackich: Warszawa 1986, 696 p. ISBN 83-04-01718-0.

Wierzbicka, A., 1997: *Understanding Cultures Through Their Key Words. English, Russian, Polish, German, and Japanese.* Oxford University Press: Oxford 1997, 317 p. ISBN 0-19-508835-2.

Wittgenstein, L., 2010: *Tractatus Logico-Philosophicus*, transl. C.K. Ogden. Project Gutenberg, http://www.gutenberg.org/ebooks/5740 (accessed 5 August 2017).

ACKNOWLEDGEMENT

Note: I declare that this manuscript (to be published in the Monograph) is my own work and it has not been published yet.