Formation of an art concept: A case study using quantitative analysis of a contemporary artist's interview data

Kikuko Takagi (qq116213@iii.u-tokyo.ac.jp)
Graduate School of Interdisciplinary Information Studies, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

Akihiro Kawase (a_kawase@ninjal.ac.jp)
Center for Corpus Development, National Institute for Japanese Language and Linguistics,
10-2 Midori-cho, Tachikawa City, Tokyo, Japan

Sawako Yokochi (yokochi-sawako@tokyomirai.ac.jp)
Department of Child Psychology, Tokyo Future University,
34-12 Senju Akebonocho, Adachi-ku, Tokyo, Japan

Takeshi Okada (okadatak@p.u-tokyo.ac.jp)
Graduate School of Education, Interfaculty Initiative in Information Studies, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

Abstract

The process of formation by an artist of an art concept for the production of a new series of artwork has not yet been empirically elucidated. The goal of this study is to describe the process of art concept formation by a contemporary artist through quantitative analyses of a text corpus based on interviews with the artist. From an analysis of the frequency of occurrence of items of vocabulary in the interview data and the TF-IDF (term frequency–inverse document frequency), we find that the second of three phases in the artist's creative process was the most critical for the formation of the art concept, as also shown in our previous qualitative study. Further, based on an analysis of co-occurrence frequencies of words, the structure of the art concept is deduced from the importance of co-occurring vocabulary. By means of visualizing the network of co-occurrence analysis, it is clarified that the feature words "The Large Glass" functioned in the first phase as the medium for dividing the structure of the concept into two parts. In the second phase, these two parts of the structure became integrated into one. In the last phase, the structure of the concept was elaborated on with the revived feature words, "White Noise" and "Duchamp".

Keywords: artistic creation, art concept, contemporary artist, quantitative analyses, interview data, case study

1. Introduction

In the field of contemporary art, the artistic creation process often results in the formation of a new art concept. Such an art concept plays an important role in contemporary art (e.g., Godfrey, 1998). This study aims to capture the processes of formation of an art concept quantitatively.

By describing the change in drawings during the creation of Picasso's Guernica, Weisberg (1986, 2006) claimed that creative problem-solving in the domain of art progresses gradually with ordinary thought processes. Through a field study, Yokochi and Okada (2005) also discovered that the images of artworks were gradually formed in the creative process of a Chinese ink painter. By analyzing the changes in the visual characteristics of artifacts in the creation process of Cubism by Picasso and Braque, Stokes (2005) proposed that paired constraints, which preclude successful responses and promote novel surprising ones, facilitate artistic creation. Through an interview study, Mace & Ward (2002) showed there are four stages in the creative activities of artists, and their processes are completed by moving back and forth between these stages. From these studies, it is clear that an artwork is created gradually through activities that change some of the features of the existing work, externalize new ideas, or collect information from the external environment, rather than that a new idea or an artwork emerges instantly. However, the detailed process of how expert artists gradually generate a new art concept in a real art creation setting needs to be studied more thoroughly. The process of formation of art concepts can be captured by analyzing interviews in which artists talk about their own creative processes (e.g., Mace & Ward, 2002; Okada, Yokochi, Ishibashi & Ueda, 2009). Okada et al. focused on the long-term span of art making activities, using retrospective interview data with artists. Though this case study pointed out the importance of analogical process in artistic creation, the detailed process of formation of a new art concept was still not clearly revealed. Takagi, Okada and Yokochi (2013) interviewed a contemporary artist every three weeks for about ten months, asking him to describe art making activities using his idea sketches, and photographs that he had taken in the preceding three weeks of work. Takagi et al. used qualitative analyses of a part of the interview data and uncovered some important aspects such as constraint modification and the use of analogy in the generation process of art concepts. However, since the data that they analyzed covered only a part of the ten-months process, a study of the whole ten-month process is still needed to capture the entire process of concept formation.

Quantitative analyses such as text mining of massive language data acquired from interviews can shed new light on the concept formation process. Some studies have used quantitative analyses of language data, such as biographies of creators, including Picasso, to uncover various factors involved in creators' success (e.g., Simonton, 2010). Basov & Nenko (2013) also presented a case study using semantic network analysis to focus on the knowledge structures underlying the creative practices of artistic communities in Spain. However, the former used rather coarse biography data and the latter was based on short-term interviews with multiple artists. In order to fully capture the concept formation process of individual artists in real art creation, we need further study focusing on a finer grain size of process data. The methods and purpose of quantitative analysis are diverse (Hearst, 1999; Feldman & Sanger, 2007). By using the quantitative analyses of a corpus of the text data acquired from interviews with an artist (Takagi et al., 2013), our study aims to reveal the change in concept through the entire process of artistic concept formation.

2. Method

2.1 Outline of the case study

The participant in this study is a male contemporary artist, Takeshi Shinohara, who studied art in Japan and in the USA and has been actively producing highly reputed artworks. He was in his late 50s at the time of the data collection. Interviews were conducted 13 times from December 2007 to September 2008 at a frequency of approximately once per three weeks (see Table 1).

The following cooperation was requested at the initial interview: (1) to create new artworks, using ideas inspired by the replica of Marcel Duchamp's *The Large Glass* (*The Bride Stripped Bare by Her Bachelors, Even*) (1923), in Komaba Museum at the University of Tokyo, (2) to take as many notes as possible of the ideas or images during the creation process, and to take photographs of hints of ideas, (3) to give regular interviews. In the first interview, he was asked about his impressions when he saw *The Large Glass*. In the second and subsequent interviews, he was asked to explain the work that had been produced since the previous interview. The questions related to, e.g.: descriptions of the contents and intentions of each drawing, sketch, note or photo; descriptions of work in production; descriptions of ideas and thoughts about the art project at the time; future plans (Takagi, et al., 2013).

2.2 Outline of generating an art concept

Shinohara generated an art concept, *White Noise* (detailed in Takagi et al., 2013), which served as the core concept of his new series of artworks, and on which he based numerous artworks. During the period of this research, in order to develop his art concept, he first drew his ideas on paper (Drawing phase) and then took photographs to collect visual information (Photography phase). After these two phases, he began the hands-on creation of the artworks (Hands-on phase). By including the Photography phase, he discovered the core part of his new art concept.

2.3 Procedure of quantitative analysis

Shinohara's utterances in all 13 interviews were divided into text units that were separated by the turn-taking of the speakers. Table 1 shows the number of text units in each phase. This study includes a co-occurrence analysis and discussion of vocabulary frequency as a quantitative analysis. More specifically, the study was conducted through a morphological analysis of the scripted protocol data. This was then followed by a word frequency analysis, co-occurrence analysis, and network centrality analysis of the extracted lexical terms. In the morphological analysis, the MeCab version 0.993 (Kudo, 2012) was adopted, and the analysis dictionary employed was the modern language version UniDic2.0.1 (Den, Ogiso, Ogura, Yamada, Minematsu, Uchimoto, Koiso, 2007). In addition, before the analysis, words related to the art concept *White Noise* ("White Noise", "Duchamp", "The Large Glass") and its subordinate concept ("word balloon") were registered manually for the analysis dictionary.

2.3.1 Word frequency analysis Firstly, in order to observe the time series variation of the art concept, which corresponds to artistic creation, we analyzed the feature words in all the interviews and in the three phases, in descending order of frequency. These feature words are the keywords in the art concept formation, namely: (1) "White Noise", the name of the art concept itself; (2) "The Large Glass", the requested theme for the new artworks, as explained in section 2.1; (3) Duchamp, the artist of *The Large Glass*. Additionally, we quantitatively analyzed feature words of subordinate concepts of "White Noise", shown in Table 2 (Takagi et al., 2013). As an analytical procedure, we compared the frequencies of word class occurrences in each interview and phase, focusing on nouns, the substantives that function as the name of a specific thing, and then extracted frequently appearing words by

Table 1: Number of text units and file size in each interview

Interview No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Phase	Drawing		Photography		Hands-on								
Number of text units	160	346	280	282	185	143	76	270	124	83	255	67	13
File size (MB)	0.68	1.40	1.20	1.20	0.76	0.59	0.31	1.10	0.53	0.34	1.00	0.28	0.07

calculating the TF-IDF index within all speech data. TF-IDF, the product of TF (term frequency) and IDF (inverse document frequency), is a weighted statistics that is used to list keywords in information retrieval and corpus analysis.

2.3.2 Network centrality analysis Based on the results of the co-occurrence analysis, we described a co-occurrence network of the top 30 terms in each phase, and applied network analysis techniques in order to numerically grasp the differences among the networks. More specifically, we used several networking algorithms based on the concept of centrality. Centrality refers to indicators that identify the most important nodes within a network. Here, we calculated three measures, which have been used in a number of applications: degree centrality, flow centrality, and Bonacich's centrality. Degree centrality is simply the number of edges of the nodes. Flow centrality measures the betweenness of the overall network or the broker role of the structure. Bonacich's centrality is a calculation based on the eigenvectors, in which a node's centrality is weighted by the centralities of its connections to other nodes (Kawase, Murai & Tokosumi, 2009).

3. Results and discussion

3.1 Word frequency analysis

Firstly, we analyzed which word class was the most basic attribute of Shinohara's utterances. Then five types of utterances, representing nouns (including pronouns), verbs, adjectives, adverbs and auxiliary verbs in Japanese, were extracted and the frequency of each was counted. Subsequently, we analyzed the frequency and the ratio of the feature words in each phase. In the word frequency analysis for the three feature terms "Duchamp", "The Large glass", and "White Noise", it appeared that the frequency of "Duchamp" and "The Large Glass" were higher in the Drawing phase, and declined in the Photography phase (see Table 3). "White Noise" did not appear in the Drawing phase, but it appeared at a rate exceeding the other two terms early in the Photography phase. In the Hands-on phase, the frequency of "White Noise" increased further. "Duchamp" occurred more frequently than "White Noise". These results suggest that, firstly, the artist, who had been inspired by *The Large Glass* by Duchamp, thought about it in the Drawing phase. Having formed his new art concept, *White Noise*, in the Photography phase, he further proceeded to elaborate on his new art concept in the Hands-on phase. Thus, the quantitative analysis shows that the Photography phase was the most important phase in generating the new art concept. This finding is consistent with the results in our previous research (Takagi et al., 2013). An analysis of the frequencies of the subordinate concepts of "White Noise" in each phase reveals that the numbers of occurrences of all four middle categories were

Table 2: Keywords for subordinate concepts of "White Noise" (excerpt)

Super-ordinate	Middle	Subordinate	Keywords				
Border	Characteristics of border	Boundary	boundary over there	border front/back	barrier up/down	wall inside/outside	fence left/right
		Net	lattice	stitch	paling	net	grid/square
		Slit/Gap Keyhole	slit	keyhole			
	Function of border	Passage	pass	drop	pass through	see through	
		Caught/ Protruding	fix	pushing out	protrusion		
		Noise/ Signal	noise	gravel	frosted glass	filter	sandstorm（on TV）
		Light	light	penetrate	reflection	shadow	transparent
Object	Relationship between objects	Medium	medium	relation	intervention	midway/ between	
		Pull	pulling each other	rice cake	word balloon	gravitation	surface tension
		White hole Black hole	white hole	black hole	entrance	exit	overflow/draining
	Instability of objects	Ambiguity	virtual image	ambiguity	blur	unclear	
		Flotation	float	air			
		Transition in time span	collapse	disappear	fall	after-image phenomenon	transfiguration

elevated in the Photography phase (see Tables 4 and 2, for specific categories (Takagi et al., 2013)). The increasing frequency of subordinate categories suggests that the attributes of visual features that were used in searching for the art concept in the Drawing phase were continually used in the Photography phase to form the new art concept, "White Noise", which is a higher order concept in making the artwork. It is also critical that the subordinate concepts of "White Noise" had already been generated in the artist's early process of concept search, before "White Noise" was discovered.

In order to extract frequent word occurrences, all the interviews were divided into 30 instances of turn-taking for each file, which created a total of 108 documents (mean capacity 178.85 KB, standard deviation 53.6). From these 108 documents, using TF-IDF, we aggregated the top 25 items of vocabulary from each of the documents. The top 30 of these terms are listed in Table 5. Table 5 shows that the three feature terms, "Duchamp", "White Noise" and "The Large Glass", ranked 7th, 15th, 22nd respectively. This means that the artist talked about these terms throughout the entire creation process.

3.2 Network centrality analysis

In order to explore the transition of the structure of the key concepts in each phase, we describe a co-occurrence network (Fig. 1) of the top 30 terms in each phase based on the results of the co-occurrence analysis. In order to reflect the influences of the frequent vocabulary network, the lengths of the edges (links) are drawn in inverse proportion to the co-occurrence frequency. In the Drawing phase (see Fig. 1), two morphological structures are depicted, centered on "image" (ranked 3rd in degree, 8th in flow, 3rd in Bonacich's centrality) and "Duchamp" (2nd, 14th, 2nd), respectively. "The Large Glass" (8th, 3rd, 8th) connects the two structures. In addition, in the Photography phase, the morphological structure seems to be embedded in one structure centered on "image" (4th, 3rd, 4th), "relation" (2nd, 4th, 2nd), and "meaning" (3rd, 2nd, 3rd). In the Hands-on phase, "Duchamp" (11th, 10th, 11th) reappears in a structure centered on "image" (2nd, 2nd, 2nd), and it is characteristic

Table 3: Frequency, proportion (%), and rank of the feature words in each phase

Phase	Duchamp	The Large Glass	White Noise
Drawing	176 (1.39), 7	176 (1.39), 7	0 (0.00), -
Photo	24 (0.21), 68	6 (0.05), 174	36 (0.33), 40
Hands-on	149 (0.51), 19	59 (0.20), 63	116 (0.40), 27

Table 4: Frequency of the middle category of subordinate concepts of "White Noise" in each phase

Middle category	Characteristics of border		Function of border		Relationship between objects		Instability of objects	
Drawing	97	(0.76%)	113	(0.89%)	66	(0.52%)	14	(0.11%)
Photography	213	(1.89%)	173	(1.54%)	145	(1.29%)	27	(0.24%)
Hands-on	268	(0.92%)	225	(0.77%)	208	(0.71%)	65	(0.22%)

Table 5: Frequency of occurrence in all interviews (top 30 terms), excluding common words

Rank	Term	Number of documents / 108	Frequency	Rank	Term	Number of documents / 108	Frequency
1	consciousness	58	1244	16	real/fact	18	213
2	man	45	413	17	spirit/mind	17	232
3	feeling	41	348	18	real	17	156
4	self	40	322	19	shadow	16	147
5	image	37	641	20	relation	16	370
6	form	34	291	21	head	16	227
7	Duchamp	32	349	22	The Large Glass	15	162
8	representation	27	511	23	stairs	14	130
9	meaning	25	590	24	completion/finish	14	246
10	edge	25	246	25	plain	14	128
11	condition	21	233	26	back	13	180
12	inside	21	220	27	light	13	166
13	front	20	190	28	dimension	13	151
14	wall	20	197	29	color	13	164
15	White Noise	18	153	30	first	12	142

that "White Noise" (24th, 24th, 24th) appears co-occurring with "meaning" (3rd, 3rd, 3rd), "image", and "representation" (4th, 4th, 4th). In the Drawing phase, the upper level of vocabulary shows that "consciousness" (1st, 13th, 1st), "Duchamp", "image", "representation" (4th, 1st, 4th) and "art" (5th, 2nd, 5th) are central concepts. Similarly, in the Photography phase, "consciousness" (1st, 1st, 1st), "relation", "meaning", "image", and "representation" (6th, 6th, 6th) appear in the upper level. "Consciousness" (1st, 1st, 1st), "image", "meaning", "representation" (4th, 4th, 4th), and "self" (7th, 7th, 7th) are the central concepts in the Hands-on phase. Among these items of vocabulary, "consciousness", "image", and "representation" are common throughout the phases in descending order of value of network centrality. However, with the higher-order terms, different items of vocabulary are central to the network in each phase. These are "Duchamp" and "The Large Glass" in the Drawing phase, "relation", "photograph" (5th, 5th, 5th) and "shadow" (8th, 8th, 8th) in the Photography phase, and "feeling" (6th, 6th, 6th) and "condition" (8th, 8th, 8th) in the Hands-on phase. These results show that the artist started to explore his new art concept from the external objects "Duchamp" and "The Large Glass", associating these with "art" and "image" in the Drawing phase. The subsequent Photography phase suggests that the artist connected "meaning" and "relation" to "shadow", which is one of the subordinate concepts of "White Noise", using "photograph". In the Hands-on phase, the results show that the artist worked on the internal representation through terms such as "image" and "feeling", while being objectively aware of his own cognitive state, taking into account the influence of the environment.

Thus, by means of visualizing the network of co-occurrence analysis, it is clarified that the feature words "The Large Glass" functioned in the first phase as the medium for dividing the structure of the concept into two parts. In the second phase, these two parts of the structure became integrated into one. In the last phase, the structure of the concept was elaborated on with the revived feature words, "White Noise" and "Duchamp".

4. General discussion

In this study, we have applied quantitative analysis to interview data on the formation of an art concept. The visualization of our analysis suggests that the process of formation of the art concept was influenced by feedback from the externalization of his ideas and by actual making of his artworks. The network centrality analysis suggests that the central words of the network changed as the artistic activity progressed and the relationship between keywords in the concepts gradually changed. These results suggest that quantitative analyses are useful methods for understanding the formation process of art concepts. We also conducted egocentric network analysis of the art concept *White Noise* with each set of interview data, focusing on the detailed changes of meaning involved in the art concept. However, this will have to be presented in a future publication due to the limitations of space in the current work.

Previous studies of artistic creation have suggested that art concept formation is influenced by drawing and by using external information to develop artistic ideas (Mace & Ward, 2002). The major contribution of our study is to quantitatively demonstrate that such processes are involved in the formation of an art concept. Because our continuous ten-months interviews with an artist provide a very rich data set, quantitative analysis was a useful measure for an objective and comprehensive understanding of the process of artistic creation.

The findings of this study have a certain consistency with the results of the studies in the domain of art, as mentioned above, and also in the domain of engineering and design. Examples of this are the analysis of the protocols and sketches produced by architects focusing on the role of the externalization of ideas (Suwa & Tversky, 1997), and the study of the design process on a design course, presenting the co-evolution of design problems and solutions, and the use of analogy (Helms & Goel, 2012). To generalize our findings, further studies of multiple cases and/or different domains will be needed.

Acknowledgments

This study was supported by Ishibashi Foundation. We thank Takeshi Shinohara for his cooperation.

References

Basov, N., & Nenko, A. (2013). Artistic Community Knowledge Structure Revealed: A Semantic Network Analysis of 'La Escocesa', Barcelona. *Centre for German and European Studies*, 3-32.

Den, Y., Ogiso, T., Ogura, H., Yamada, A., Minematsu, N., Uchimoto, K., & Koiso, H. (2007). The Development of an Electronic Dictionary for Morphological Analysis and its Application to Japanese Corpus Linguistics. (In Japanese) *Japanese Linguistics*, 22, 101-123.

Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge: Cambridge University Press.

Godfrey, T. (1998). *Conceptual Art (Art and Ideas)*. London: Phaidon Press.

Hearst, M. A. (1999). Untangling Text Data Mining. *The 37th Annual Meeting of The Association for Computational Linguistics*, University of Maryland, June, 20-26, 1999, 3-10.

Helms, M., & Goel, A. (2012) Analogical Problem Evolution in Biologically Inspired Design. *Fifth International Conference on Design Computing and Cognition*, College Station, Texas, U.S.A.

Kawase, A., Murai, H., & Tokosumi, A. (2009). Conceptual Transitions in Papers of Music Criticism Extracting Musical Concepts based on Network Centrality. *Journal of Japanese Society of Information and Knowledge*, 19, 138-143.

Kudo, Y. (2012). MeCab: Yet Another Part-of-Speech and Morphological Analyzer, https://code.google.com/ (accessed 2015-01-22)

Mace, M. A., & Ward, T. (2002). Modeling the Creative Process: A grounded theory analysis of creativity in the domain of art making. *Creativity Research Journal,* 14, 179-192.

Okada, T., Yokochi, S., Ishibashi, K., & Ueda, K. (2009). Analogical modification in the creation of contemporary art. *Cognitive Systems Research*, 10, 189-203.

Simonton, D. K. (2010). Creativity in Highly Eminent Individuals. In J.C. Kaufman, R.J. Sternberg (Ed.), *The Cambridge Handbook of Creativity*. Cambridge: Cambridge University Press.

Stokes, P. D. (2005). *Creativity from Constraints: The Psychology of Breakthrough*. New York: Springer Publishing Company.

Suwa, M., & Tversky, B. (1997). What do architects and students perceive in their design sketches? A protocol analysis. *Design Studies*, 18, 385-403.

Takagi, K., Okada, T., & Yokochi, S. (2013). Formation of an art concept: How is visual information from photography utilized by the artist in concept formation? *Cognitive Studies*, 20 (1), 59-78.

Weisberg, R. W. (1986). *Creativity: Genius and other myths*, Oxford: Freeman.

Weisberg, R. W. (2006). *Creativity: Understanding Innovation in Problem Solving, Science, Invention, and the Arts*. New Jersey: Wiley.

Yokochi, S., & Okada, T. (2005). Creative Cognitive Process of Art Making: A Field Study of a Traditional Chinese Ink Painter. *Creativity Research Journal*, 17, 241–255.

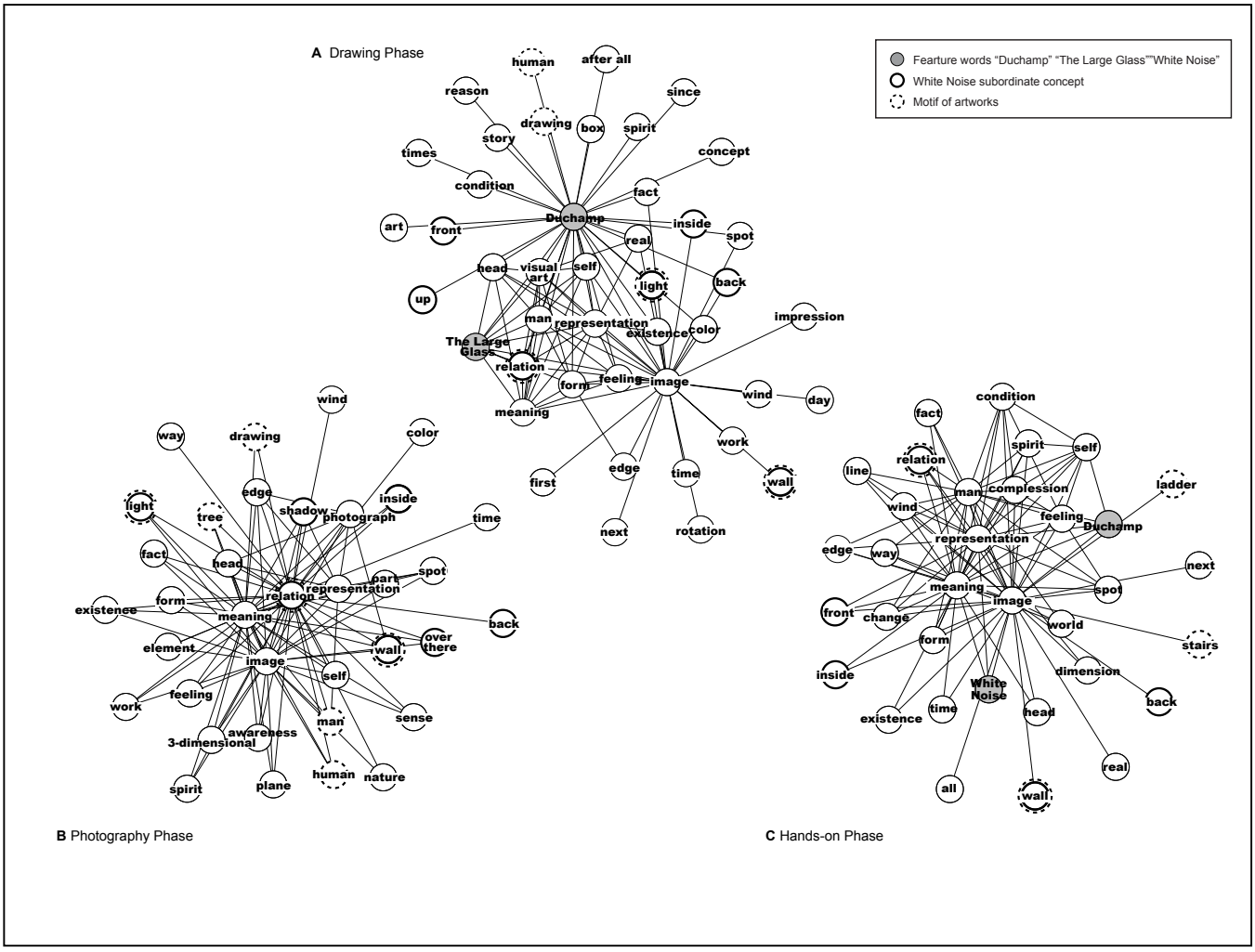

Figure 1: Network of co-occurrence in each phase (excerpts from top 30 terms)