The University of Osaka
Institutional Knowledge Archive

Title	ロシア語の語彙連想とグラフの連結成分について
Author(s)	上原, 順一
Citation	言語文化研究. 2018, 44, p. 19-31
Version Type	VoR
URL	https://doi.org/10.18910/68011
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

ロシア語の語彙連想とグラフの連結成分について

上　原　順　一

Word association in Russian and connected components of graphs

UEHARA Junichi

Summary: The purpose of this study was to explore the properties of word association networks（graphs）in Russian. In the First section I created an ideal tree diagram which expresses the processes of word association ― a stimulus word has tree response words, with these tree words set as stimuli and procedures repeated. In the Second section I analyzed the graphs of 10 high frequency words as stimuli words. In this study main stress falls on behavior of words（vertices）whose betweenness centrality is 0 and in-degree is two or more. Though these kinds of typically appear to plays minor roles in connecting vertices, a lot of them in the graphs mentioned above are shared with the smaller graphs and link them as connected components.

キーワード：ロシア語，連想，グラフ

0．研究の発端

ロシア語の連想辞典を用いれば，連想ゲームのような遊びができる。たとえば，белый「白い」から連想される語のひとつにкирпич「通行止めの道路標識，煉瓦」がある。同様に，кирпич「煉瓦」から連想される語にкаратэ「空手」がある。さらに，каратэ「空手」→Япония「日本」と続けられる。擬人的に表現すれば，「白い」さんの友達の友達の友達が「日本」さんであると言える。また，белый から連想されるのはкирпич以外にもある。たとえば，снег「雪」，цвет「色」である。さらにそれらから連想される語もそれも連想ゲームに加えると，このбелый「白い」から始まる連想の体系（連想野）は，かなり多くの語から成り立つ樹形図のような形をとるはずである。

多数の語から成り立つ連想野における語と語の近さ，たとえれば，人と人が意外に近いところで知り合いかもしれないことは，スモールワールドと呼ばれている。筆者は，かつて，ロシア語の高頻度語をとりあげ，それから連想される語と，さらにそれらから連想される語を連想

辞典で調べた結果，語や連想関係の総数が数千にもおよぶ連想ネットワークを求めることができた。その中では，語から他の語にとどく連想は，平均的には 3 ステップであることを述べた［上原 2016］。

同時に，スモールワールド性はウェブサイトのリンク，会社の持ち株関係，電力網など，あらゆるネットワークのもつ性質であると述べる，グラフ理論の先行研究も紹介した。つまり，筆者が調べた連想連鎖が，よくあるネットワークのひとつであることを，確認した。ただ，「確認した」以上の議論を進めることはできなかった。それは，連想におけるこの語の近さがどのようにして生じているかについて，議論が足りなかったと言うことである。今回の論文では，語と語の近さは，どのような仕組みで成り立つのか，ないしは，この仕組みはどのような手続きで求めることができるのか，述べたいと思う。とくに，個別に存在する連想野をつなげる語があれば，それはスモールワールドを実現している重要な因子であると考えられるので，注目することになる。

今回の論文では，これを求めるひとつの過程を考えたい。なお，用語としては「ネットワーク」は「グラフ」に統一した。

1．連想とは

「連想」なる用語の定義はいくつかあるが，これを刺激語から反応語へのつながりとする見方は一般的である。たとえば，『新編 認知言語学キーワード事典』の「連想」の項目には，「ある語（句）が与えられたとき，それと関連する他の語（句）あるいは事柄が思い浮かぶ心的作用」とあり，「最初に与えられる語を刺激語，引き出される語を反応語と呼ぶ」と続く［辻（編）2013: 375］。ロシア語の文献でも，同じような記載が少なからず見られる。たとえば，『認知言語学ハンドブック』は，「連想」が「二つの現象，イメージ，モノなど―通常は刺激とそれに付随する反応―を結びつけること」［Кубрякова Е.С., и др 1996: 13］であると記している。

連想が研究の素材となっている資料としては，各種の連想辞典がある。これらは多くの人びとが利用でき，また，取り扱う手順が同じならば，想定される研究結果も同じになる客観性が大きな利点である。ロシア語の連想辞典をひもとけば，おおむね，次のような記載がある。

まず，ソ連・ロシアの言語学者レオンチエフが出版した連想辞典『ロシア語連想規範事典』［Леонтьев 1977］をとりあげる。これは，ソ連で出版されたロシア語の連想辞典でパイオニア的な存在である[1)]。これで город「都会，市」の項目を見よう。

город「都会，市」― большой「大きな」29, деревня「田舎，村」27, село「村」21 … школа

1）なお，連想辞典の編纂史は，レオンチエフの連想辞典や，ウフィムツェワらの論文［Уфимцева, Черкасова 2004］がまとめている。

「学校」1. N=237.

これはгородなる刺激語に対して，большой「大きな」なる反応語を返した被験者が29人であること，被験者の総数が237人であることなどを表している。

別の連想辞典，カラウーロフらが著した『ロシア語連想辞典』には上と同じ刺激語について次のような記載がある［Караулов Ю. Н. и др. 1994］。

город「都会，市」— большой「大きな」55, деревня「田舎，村」36, герой「英雄」24…
юности「青年時代の」1; 549 + 213 + 1 + 151

反応語を返した被験者数の表示は上の辞典と書き方が同じである。一方，末尾に記載されている数字がこの辞典の特徴である。第 1 の数字は反応総数，第 2 の数字は反応語の種類がいくつであったかを示す。第 3 の数字は反応語を返さなかった被験者の数，4 番目の数字は 1 人の被験者のみが返した反応語の数である。

これらの情報はグラフ理論で理解することが可能である。刺激語と反応語はグラフ理論における頂点，連想関係は辺であると考えられる。連想関係には方向があるので，有向グラフとなる。図示すれば下のようになる。反応語ごとの反応数あるいは反応数を反応総数で割った値は辺の属性として設定することが可能であるが，ここでは省略した。

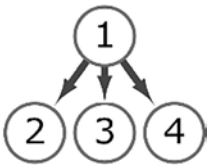

図 1　連想辞典のデータをグラフ化したもの

レオンチエフの記述をあてはめれば，（1）がгород「都会，市」，（2）がбольшой「大きな」，（3）がдеревня「田舎，村」，（4）がсело「村」となる。

一方，「連想」には，連想ゲームのような連想の連鎖とでも名付けられる現象がある。これは，ある刺激語と反応語を得たあとに，その反応語を新たな刺激語として反応語を得ること，ないしはその繰り返しである。当初の刺激語が「山」なら，その反応語「川」がありうる。その「川」を刺激語に設定すれば，たとえば「水」が反応語として返されるだろう。この連想の連鎖は，刺激語と反応語のペアを記述する連想辞典から得ることができる。刺激語を 3 語，その反応語を 3 語だけ，さらにその反応語を刺激語として新たな反応語を得る過程は，次のように図解できる。

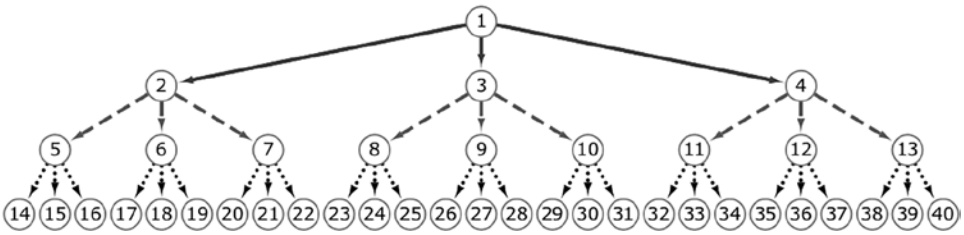

図２　連想野の例

この図に説明を加える。実線は，当初の刺激語（1）から得られる反応語が（2～4）であることを，破線は先に得られた反応語（2～4）を刺激語とすると，反応語（5～7，8～10，11～13）が求められることを，点線は先に得られた反応語，たとえば（5）を刺激語にすると反応語（14～16）が得られることを示している。このようにして得られた反応語の集合は，連想野と呼ぶことがある［Кобозева 2000: 108, Тарасов Е.Ф. и др. 2017: 9］。刺激語と反応語のペアは，語の連想関係としては，いわば最小単位として取り扱えるものであると考えられる。

このような連想野では，たとえば（1→2）を第１レベルの連想関係，（2→5）を第２レベルの，（5→14）を第３レベルの連想関係と名付けることができる。また，第１レベルと第２レベルの連想関係をまとめて「第２レベルまでの連想関係」と言うことも可能である。

ある刺激語から得られる反応語の数は多い。上で見たように『ロシア語連想辞典』は，город「都会，市」の反応語が（種類としては）213あると書いている。これらの反応語は一様に扱うこともできる。これは，語を刺激語との連想関係があるかないかで区別する方法である。一方，反応数が多い語を重視するのが良いという意見もある。『ロシア語学習連想事典』は，刺激語に対する反応語のうち高頻度の語のみを選んで教材にしているが，この著者は「高頻度の反応語は母語話者大多数の意識に保持されているものであり，言語意識の基準を成立させている」と述べている［Тарасов Е.Ф. и др. 2017: 9］。この議論によれば，高頻度の反応語のみ扱うことは，連想の研究にとっては有意義であろうと考えられる。筆者は，刺激語ごとに得られる反応語は３語ずつ扱うことにした。

理論的に得られる連想野の特徴をグラフとして観察しよう。いわゆる根は当初与えられた語のみとなる。根以外の頂点数は，第１レベルが３，第２レベルが９，第３レベルが27となる。根を含めると，この連想野グラフには40の頂点がある。これは，上の図２を参照されたい。当初の刺激語を10語にすると，当然のことながら，この図のようなグラフが10得られる。根の数は10，それ以外の頂点数は第１レベルが30，第２レベルが90，第３レベルが270となり，すべての頂点数は400となる。

ただ，このような連想野と，現実的に得られる連想野は，かなり異なる。以下，理論的な連想野と実際の連想野との比較をもとに，それらの異なり具合，更に言えば，理想的な連想野としての樹形図がいかに変化することで実際に得られる連想野が成り立つのかを観察したい。

2．実際の連想野

拙論では，形容詞，名詞，動詞の高頻度語から10語を選んだ。それらを刺激語として得た反応語のうち，反応数が多い順に 3 語を抽出した。その 3 語をさらに刺激語として，繰り返し反応語を求めた。第 3 レベルまでこの作業を行い，連想野を得た。それらの連想野をここでは「形容詞グラフ」,「名詞のグラフ」などと呼ぶ。また，それぞれに，第 1 レベルから第 3 レベルまでの連想関係を表したグラフを作成することが可能である。これは,「第 3 レベルまでの名詞グラフ」などと名付ける。詳細は本論末尾の「作業の手順」における［手順番号 : 1］を参照されたい。

次頁の図 3 は第 3 レベルまでの名詞グラフを図示したものである。

当初の刺激語が10あったので，10の連想野が想定されるが，実際には 6 しかない。図 2 の樹形図が，理論的には10だけばらばらにあったのだが，ある程度までまとまったと考えられる。

これは樹形図が形を変えた，あるいは壊れたとも見える。この様子をグラフ理論の概念で表現すると次のようになる。

グラフには「連結」,「非連結」の 2 種類がある。「グラフのどの 2 頂点についても，その 2 頂点を結ぶ歩道が存在するとき，そのグラフは連結であるという」。そうでないグラフは非連結である。さらに，非連結グラフはいくつかの連結グラフに分かれることがあり，この連結グラフは，非連結グラフの連結成分と呼ばれている［小林 2013: 25］。

図 3 のグラフは，全体的には非連結グラフである。また，これには 6 つの連結成分がある。単純に表現すれば，6 つの島から列島がなりなっているようなものである。

連結成分のあり方は，形容詞グラフ，名詞グラフ，動詞グラフとレベルによってさまざまである。

表 1　連結成分の数

	形容詞グラフ	名詞グラフ	動詞グラフ
第 1 レベルまで	5	10	5
第 2 レベルまで	3	8	2
第 3 レベルまで	1	6	1

先に示したように，第 3 レベルまでのグラフでは，名詞グラフのみが非連結グラフで，6 つの連結成分をもっている。第 3 レベルの形容詞グラフ，動詞グラフは連結成分が 1 つずつであり，両グラフは連結グラフである。これら 2 つのグラフは，刺激語と反応語がなんらかの形で，つながっていることになる。

第 3 レベルまでの名詞グラフの連結成分のうち，頂点数が最大のものは図 3 の上半分に位置

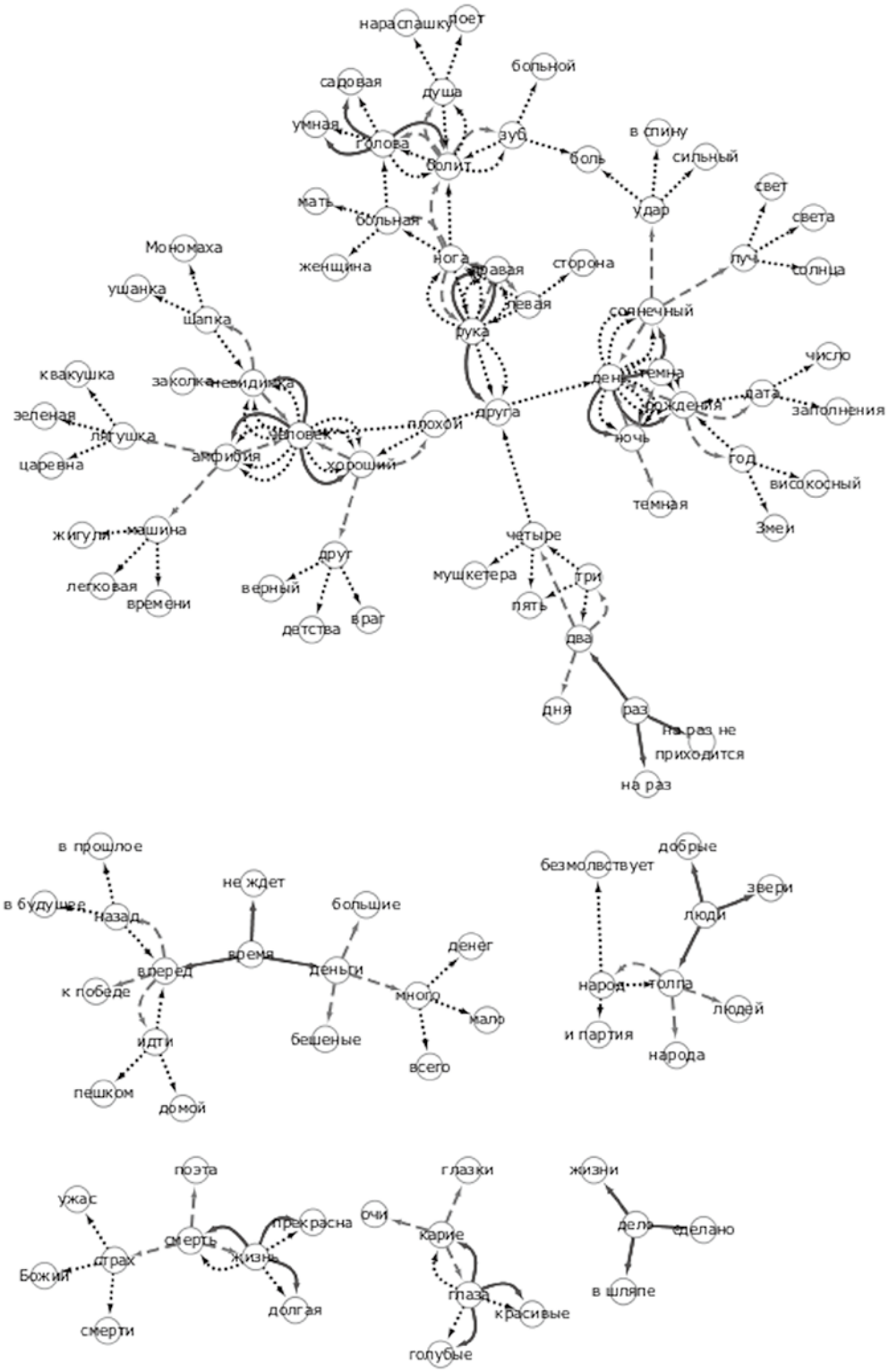

図３　第３レベルまでの名詞グラフ

する。この当初の刺激語は，голова「頭」，рука「手」，раз「回」，день「日」，человек「人間」であるグラフである。他の 5 つの連結成分がもつ当初の刺激語はそれぞれ，время「時」，жизнь「命」，люди「人々」，глаза「目」，дело「用事」1 語ずつである。

この連結・非連結のあり方は，原因を品詞のみに求めることは難しい。それは，いわゆる変化形の取り扱いである。たとえば，図 3 の右下にある，дело「用事」を当初の刺激語とするグラフには，第 1 レベルの反応語 сделано「行われた」，в шляпе「帽子を着用している（дело в шляпе は「物事がうまくいく」なる意味の慣用句）」，жизни「命の」しかない。このような変化形は連想辞典には登録されていないことが多く，連鎖が広がらないと考えられる。変化形がグラフの末端頂点になっているのである。

逆に，頂点にはグラフをつなぐものもあり，連結成分のあり方にとって重要であると考えられる。これを観察しよう。

第 3 レベルまでの名詞グラフにおける，頂点数が最大の連結成分にある，друга あたりをみることにする。これは друг「友人」の生格，「友人の」である。次の図を参照されたい。

図 4−1 はこのグラフの全貌，図 4−2 は друга のあたりである。図 4−2 の中央やや右に друга「友人の」がある。その上方に рука「手」がある。矢印の向きをみると，рука → друга が認められる。一方，下の方には раз「回」がある。これからたどっていくと，раз「回」→ два「2」→ четыре「4」→ друга「友人の」という連想連鎖が見える。つまり，この друга は当初の刺激語が рука であるグラフと，当初の刺激語が раз であるグラフをつないでいる。この様子は，рука →（друга）← четыре ← два ← раз と表現できる（括弧内は問題としている語）。

このようにグラフをつないでいる頂点は，当初の刺激語の数だけ連結グラフがある（グラフがばらばらに存在する）はずのグラフを，連結グラフの数が少ない実際の連想グラフにする機能を担っている。換言すれば，この頂点がなければ，グラフが分離されてしまう「切断点」になっている。

このようにグラフの連結にとって重要な頂点は，どのような条件を与えれば抽出できるだろうか。ひとつの基準は媒介中心性である。これは「そのノード（頂点）を通過しないと他のノード（頂点）に到達できない度合い，ある点がその他の 2 点を結ぶ最短経路である度合いであり，値が大きいほど中心性が高い」［金 2009: 65−67］とされている。鉄道網を例にとれば，駅から他の駅に最短で行く際に通過することが多い（おそらく都心の）駅は媒介中心性が大きく，逆に終点の駅は通過することがないので媒介中心性は 0 である。

もうひとつの基準は，入次数である。これは，ある頂点に入る，有向の辺の数である。これが 1 ならばその頂点は末端であるが，2 以上あれば，グラフと他のグラフをつないでいる可能性がある。

作業の過程で，これらの基準を同時に満たす頂点は，一定の手続きで抽出することが可能であることがわかった。（詳細は［手順番号 : 2］）を参照されたい。

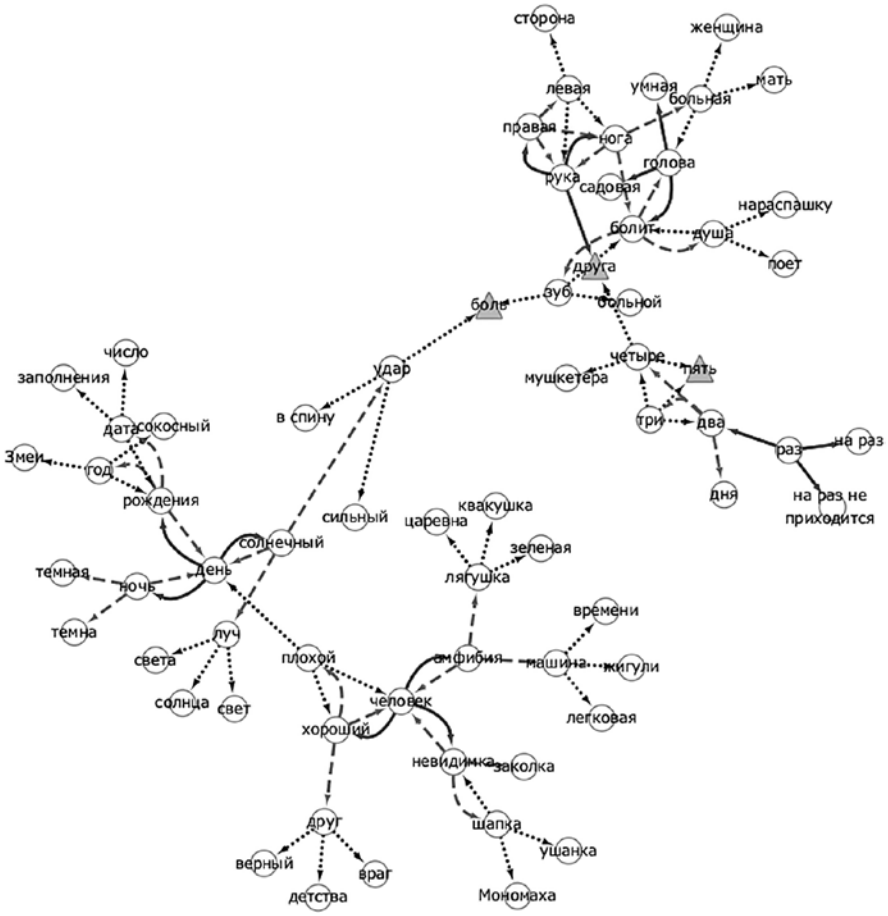

図 4-1　名詞グラフの例

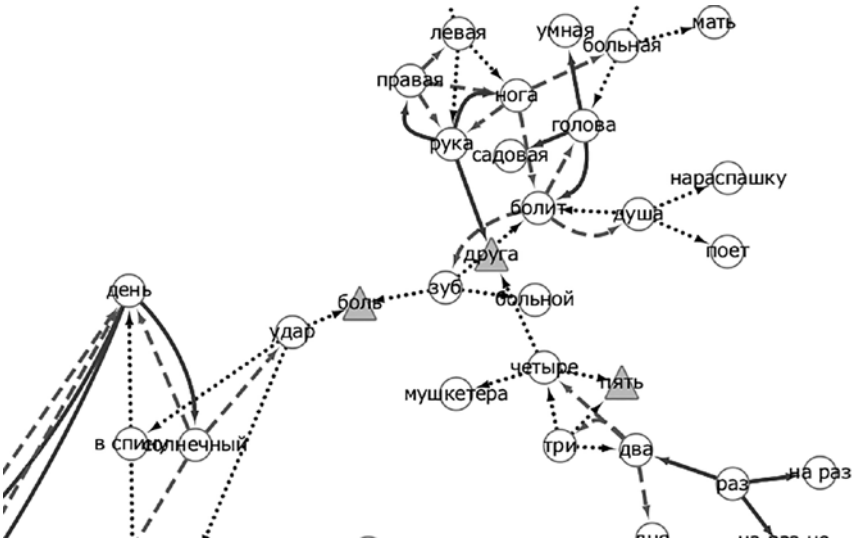

図 4-2　друга のあたり

図 4−1 と 4−2 でみた，グラフの例では，得られた頂点には，друга 以外に боль「痛み」，пять「5」がある。後者は切断点ではないが，前者は，день「日」→ солнечный「日光の」→ удар「ショック」→（боль）「痛み」← зуб「歯」← болит「痛む」← голова「頭」である。солнечный удар は「日射病」。色つきの△印が抽出された語である。

形容詞では，第 3 レベルまでのグラフは連結成分が 1 つしかないので，比較対象は第 2 レベルまでのグラフで頂点数最大の連結成分を見よう（図 5）。これからは，родной「生まれの」，мой「私の」が抽出できた。前者は，далекий「遠くの」→ край「地方」→（родной）「生まれの」← дом「家」← новый「新しい」，後者からは，маленький「小さな」→ ребенок「赤ちゃん」→（мой）「私の」← дом「家」← новый「新しい」が求められた。

第 2 レベルまでの動詞グラフで頂点数最大の連結成分からは，мочь「できる」→ делать「行う」→（дело）「用事」← слово「ことば」← сказать「言う」，видеть「見える」→ слышать「きこえる」→（хорошо）「よく」← жить「生きる」← стать「…になる」← быть「ある，いる」，знать「知っている」→（всё）「すべてのことを」← мочь「できる」，говорить「言う」→（правду）「真実を」← сказать「言う」，знать「知っている」→ много「たくさんの」→（денег）「お金」← нет「ない」← есть「ある」← хотеть「…したい，欲する」← мочь「できる」，говорить「言う」→ молчать「黙っている」→（долго）「長い間」← спать「寝ている」← хотеть「…したい，欲す

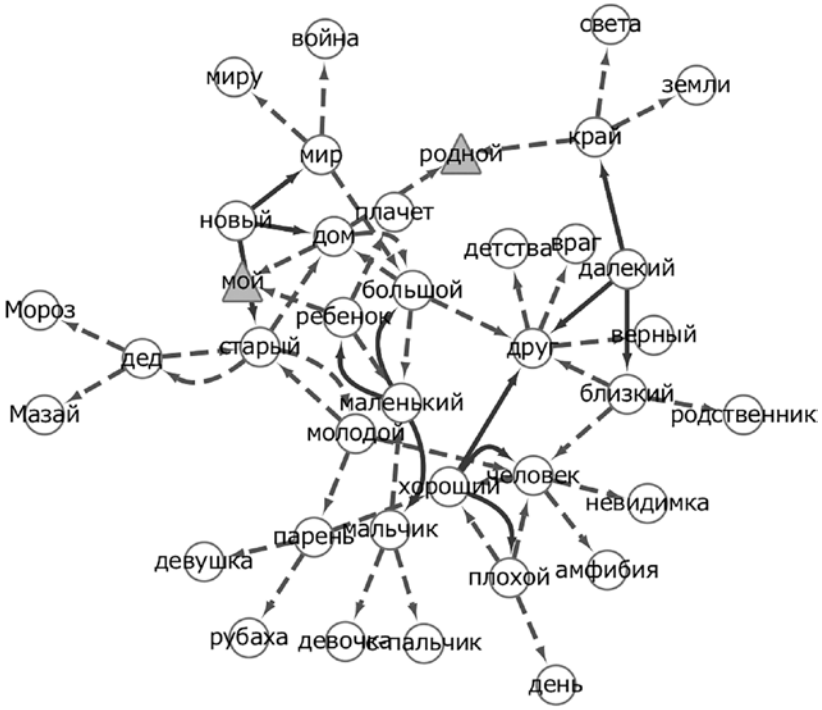

図 5　形容詞グラフの例

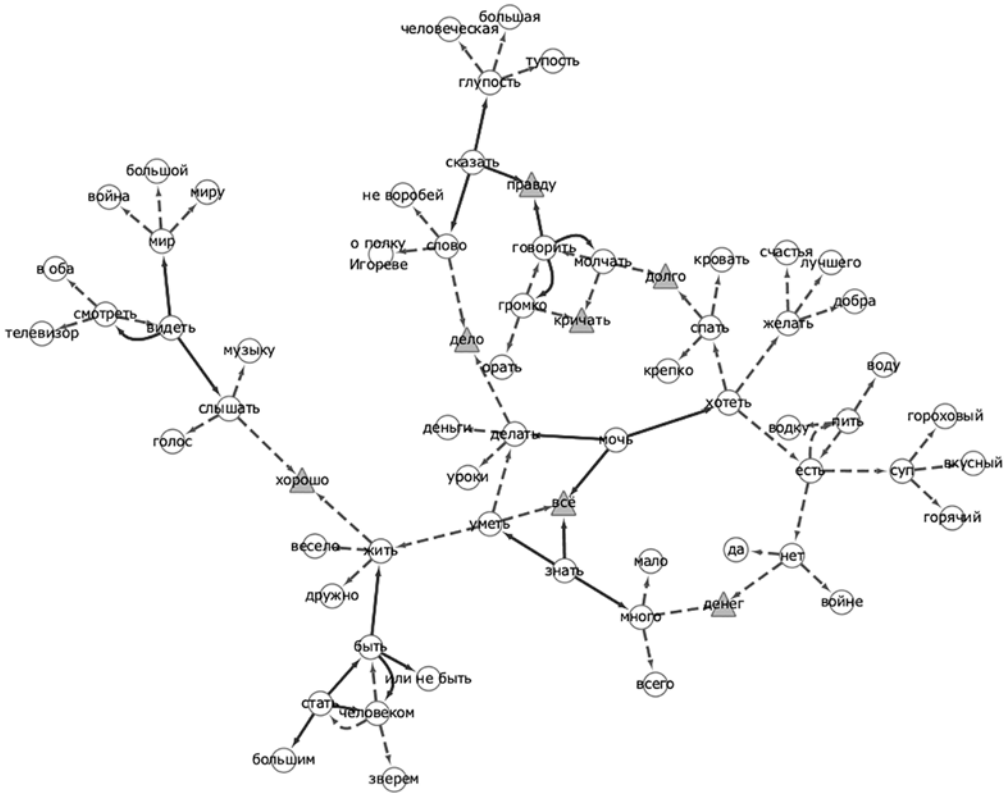

図 6　動詞グラフの例

る」← мочь「できる」が抽出できた。同じ条件で抽出できた кричать「叫ぶ」は切断点ではない（図 6）。

3．まとめ

頂点の次数が多いものは，他の頂点とのつながりが多い。頂点の媒介中心性が大きければ，それは他の頂点が連絡するのに役立っているということである。この論文で取り出したのは，逆に，次数が小さく，媒介中心性は極度に低い頂点であった。これらのすべてではないが，大多数が，より小さなグラフの連結には重要な機能をもっていることがわかった（括弧内の語）。人付き合いにたとえれば，知り合いの数が少なくても，そして，その人を介して情報はあまり伝わらなくても，集団と別の集団の橋渡し役としては十分な働きをしている人たちがいるということだろう。

今回の論文では，これらの語を詳細に分析するところまでは手が届かなかった。ただ，それらが「意外と」普通の語であることは指摘できるだろう。

当初の刺激語は高頻度語ではあるが，それから得られる反応語が高頻度であるとは限らない。

連想はあくまで連想であり，「意外な」語も散見された。たとえば，человек「人間」なる刺激語から得られる反応語（反応数では 1 位と 3 位）は，невидимка「姿を消すことができる人」とамфибия「両生類」である。前者は человек-невидимкаで「透明人間」である。私たちの前にあるのはデータのみであり，連想される過程はなぞである。しかし，これがおそらく小説のタイトルであろうことは推測できる。後者については “Человек-амфибия” というソ連映画がある。このような，いわば，ふたを開けるまでわからない反応語が見られる中では，拙論で得た，複数グラフが共有する語は，「意外と」普通であった。

筆者は先の論文で，「語の関係をグラフというメタファーをもとに調べていくと，この研究分野に新たな理論的基盤が発生するのではないだろうか」と書いた［上原 2016: 37］。今回の論文でグラフとしてのメタファーに登場したのは，次数も中心媒介性も小さな語である。率直に述べれば，理論的基盤に言及するには，グラフ理論で既知のことがらをさらに丁寧に語の連想連鎖に当てはめて考える必要があり，筆者はこの作業を行いつつある段階である。

連想野は，当初の刺激語を今回扱った動詞，形容詞，名詞だけではなく，ロシア語では自立語と考えられている副詞，数詞などにも範囲を広げて解明手続きや結果の妥当性を検証する必要があろう。

また議論をさらに進める機会があれば，今回得た「意外と普通の語」を，グラフ研究の脈絡で人間関係を表現するのによく使われる現象である，「弱い紐帯の強み」として観察することを視野に入れたい。これはつながりの弱い人から有用な情報がもたらされることがあるという人付き合いのありさまを表現している。筆者が数字や図で説明しようとしている語の関係において，「普通」や「人間関係」といった血の通った概念を持ち出すのは，筆者の説明力が不足しているからであろう。ただ，語の関係を観察していると，語が人間のように見えてくるのは，そして，グラフと人間の間にあるメタファーから逃れにくいのは，偽りのないところである。

4．作業の手順

拙稿作成のために行った作業の手順を示す。番号は本文中の番号と一致する。

利用した主なソフトウェアやウェブで得た情報源は次の通り。

Cytoscape: Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research 2003 Nov; 13(11): 2498–504 URL http://www.cytoscape.org/

［手順番号 : 1］

まず，С. А. Шаров 氏が公開していた語彙の頻度辞典より，形容詞，名詞，動詞の高頻度語を 10 語ずつ取り出した。それは，

形容詞 : хороший, новый, последний, черный, старый, белый, молодой, маленький, далекий, правый

名詞 : рука, время, человек, дело, раз, глаза, жизнь, голова, день, люди

動詞 : быть, сказать, мочь, говорить, знать, есть, хотеть, стать, видеть, идти

である。ファイルは http://dict.ruslang.ru/freq.php（2016 年 4 月ごろ閲覧）で得たもの。現行の頻度とはやや異なる。

これらをひとつひとつウェブ版の連想辞典 www.tesaurus.ru/dict/dict.php（2017 年 7 月から 9 月に閲覧）に入力して，反応語のデータを得る。これを第 3 レベルまで繰り返す。結果的に次のような表計算用テキストファイルを作成した。なお，この辞典の説明は，http://it-claim.ru/Projects/ASIS/（閲覧時期は上と同じ）の 2 つめにある。これによれば，データは 2 つの連想辞典をまとめたもので，最大のロシア語連想辞典。

コメント, No, S, 反応総数, R, 反応数, Level
#,2,,хороший,#,1
,3,хороший,642,человек,95,1
,4,хороший,642,друг,90,1
,5,хороший,642,плохой,79,1
#,131,,,человек,#,2
,132,человек,643,невидимка,25,2
,133,человек,643,хороший,25,2
,134,человек,643,амфибия,22,2
#,434,,,невидимка,#,3
,435,невидимка,105,человек,34,3
,436,невидимка,105,шапка,34,3
#,437,невидимка,105,/,3,3
,438,невидимка,105,заколка,2,3

ヘッダ部分の「コメント」は # がある行は Cytoscape に読み込まない設定にできるので，適時利用する。ここでは刺激語の覚え書きにした。“No”は整理番号。“S”は刺激語，「反応総数」は今回は利用しなかったが，「反応数」とともに用いることで，連想関係である重みの属性とすることができる。「反応数」はその刺激語を回答した人数。拙論では 3 位までを連想野に加えた。“R”は反応語のこと。“Level”は連想連鎖のレベルを示している。なお，“R”の欄にある“/”は反応語の回答がなかったことを著している。上の例では，下から 2 つめに相当する。この場合は，下位の反応語を繰り上げた。

このようなテキストファイルをCytoscapeに読み込むことが可能である。今回の作業では，オプションとして，Delimiterにコンマやタブを指定し，1行目は列のタイトルとして使う，#で始まる行は無視する（コメント行として使う）ことにした。また，「刺激語」を“Sourse Node”，「反応語」を“Target Node”に割り当てた。

なお，実際の作業では，上のテキストファイルは，便宜上，当初の刺激語の品詞ごとに，そしてレベルごとに設定した。

［手順番号: 2］

Cytoscapeのコントロールパネルでフィルタを設定することで切断点をある程度まで絞り込んで抽出することができた。フィルタは本文にあるように，「媒介中心性 Betweenness Centrality が0」，「入次数 Degree In が2以上」の2条件が同時に成立するようにする。また，入次数のフィルタが厳格に適用されるように，重複する辺はあらかじめ削除した。

参考文献

Ахманова О.С., Словарь лингвистических терминов, Москва, 1966.

Караулов Ю.Н. и др., Русский ассоциативный словарь, Книга 1−2, Москва, 1994.（『ロシア語連想辞典』）

Кобозева И.М., Лингвистическая семантика: учебник, Москва, 2000.

Кубрякова Е.С., и др., Краткий словарь когнитивных терминов, Москва, 1996.（『認知言語学ハンドブック』）

Леонтьев А.А., Словарь ассоциативных норм русского языка, Москва, 1977.（『ロシア語連想規範辞典』）

Тарасов Е.Ф. и др., Учебный ассоциативный словарь русского языка, Санкт-Петербург, 2017.（『ロシア語学習連想辞典』）

Уфимцева Н.В., Черкасова Г.А., Ассоциативная лексикография и исследования языкового сознания // Филология и культура（No.4, 38), 2004.

上原順一，「語彙のネットワークについて―ロシア語の連想連鎖―」，『言語文化研究』第42号，pp.25−42，大阪大学言語文化研究科，2016.

金明哲，『テキストデータの統計科学入門』，岩波書店，2009.

小林みどり，『あたらしいグラフ理論入門』，牧野書店，2013.

辻幸夫（編），『新編 認知言語学キーワード事典』，研究社，2013.