

# 国立国語研究所学術情報リポジトリ

## Construction of Russian Translation Data of "The Balanced Corpus of Contemporary Written Japanese"

メタデータ	言語: jpn 出版者: 公開日: 2019-02-14 キーワード (Ja): キーワード (En): 作成者: 宮内, 拓也, Prokhorova, Maria, PROKHOROVA, Maria メールアドレス: 所属:
URL	<a href="https://doi.org/10.15084/00001633">https://doi.org/10.15084/00001633</a>

## 『現代日本語書き言葉均衡コーパス』の ロシア語翻訳データの構築

宮内 拓也 (東京外国語大学大学院 / 日本学術振興会特別研究員 / 国立国語研究所共同研究員)\*  
プロホロワ・マリア (東京外国語大学大学院)

### Construction of Russian Translation Data of “The Balanced Corpus of Contemporary Written Japanese”

Takuya Miyauchi (TUFS / JSPS / NINJAL)  
Maria Prokhorova (TUFS)

#### 要旨

『現代日本語書き言葉均衡コーパス』(の一部のデータ)には、既に英語、イタリア語、インドネシア語、中国語の翻訳データが構築されているが、新たにロシア語の翻訳データを構築した。対象となるテキストは『現代日本語書き言葉均衡コーパス』新聞 (PN) コアデータ 16 サンプル (総語数は短単位で全 16,657 語) とし、ロシア語翻訳データの総語数は 13,070 語となった。本データの構築あたっては、日本語からロシア語へ人手による翻訳を行った。また、日本語とロシア語の言語構造の違いにより、翻訳に困難を生じさせた箇所も多くあった。本稿では、翻訳データの構築方法、翻訳の際の留意点の詳細を述べる。さらに、この翻訳データの構築により、原データと並べることで疑似的な日露対訳コーパスとしての利用も可能であり、本データは日露対照研究に活用できると考えられる。本稿では、そのような活用の一例として、日本語の文末形式について、簡単にロシア語と対照させて論じる。

#### 1. はじめに

『現代日本語書き言葉均衡コーパス』(Maekawa et al. 2014; 以下, BCCWJ) (の一部のデータ)には、既に英語、イタリア語、インドネシア語、中国語の翻訳データが構築されている。今回、新たに人手による翻訳により、BCCWJ の 16 サンプル分のロシア語の翻訳データを構築した。本稿では、翻訳データの構築方法、翻訳の際の留意点や翻訳データの活用の可能性を報告する。

以下、2 節では翻訳対象としたデータと構築したデータの概要について、3 節では翻訳の方法と翻訳の際の留意点について、4 節では翻訳データの日露対照研究への活用の可能性についてそれぞれ述べる。5 節は本稿全体のまとめである。

---

\* miyauchi.takuya.k0 @ tufs.ac.jp

## 2. 翻訳対象のデータと翻訳データの概要

翻訳の対象は BCCWJ の新聞 (PN) コアデータ 16 サンプルである。サンプルの選択は BCCWJ-ANNOTATION-ORDER<sup>(1)</sup> に基づく。対象の基本的なデータとして、表 1 に対象となるテキストの総語数 (短単位数), 文節数, 文数を示す。

表 1 対象となるテキストのサイズ

サンプル名	短単位数	文節数	文数
PN1c.00001	784	236	42
PN1d.00001	783	235	34
PN1e.00001	763	219	35
PN1f.00001	797	181	38
PN2e.00001	750	214	27
PN3b.00001	975	311	43
PN3g.00001	2,640	919	142
PN4a.00001	1,244	425	51
PN4b.00001	758	246	26
PN4c.00001	737	250	31
PN4f.00001	1,047	297	40
PN4g.00001	905	296	36
PN1a.00002	1,797	611	93
PN1b.00002	1,024	277	38
PN1d.00002	734	206	28
PN1e.00002	919	272	35
計	16,657	5,195	739

構築されたロシア語翻訳データの総語数は 13,070 語であり、文数は 848 文であった。表 2 に対象とする各サンプルごとの語数, 文数を示す。

既に述べたように、『現代日本語書き言葉均衡コーパス』(の一部分のデータ)には、英語、イタリア語、インドネシア語、中国語の翻訳データがある。これらの BCCWJ 外国語翻訳データと比較すると、ロシア語が今のところの規模としては最も大きい。各言語の翻訳データの総語数<sup>(2)</sup>, 対象のサンプル数をまとめたものが表 3 である。

<sup>(1)</sup> BCCWJ コアデータサンプルにおけるアノテーション優先順序である。以下参照のこと。 <https://github.com/masayu-a/BCCWJ-ANNOTATION-ORDER>

<sup>(2)</sup> ただし、中国語は字数をカウントしている。

表2 ロシア語翻訳データのサイズ

サンプル名	語数	文数
PN1c_00001	568	40
PN1d_00001	555	54
PN1e_00001	653	34
PN1f_00001	568	56
PN2e_00001	639	33
PN3b_00001	787	39
PN3g_00001	2,227	167
PN4a_00001	1,012	62
PN4b_00001	548	23
PN4c_00001	587	33
PN4f_00001	761	38
PN4g_00001	674	36
PN1a_00002	1,409	106
PN1b_00002	763	57
PN1d_00002	564	29
PN1e_00002	755	41
計	13,070	848

### 3. 翻訳方法

翻訳データの構築あたっては、日本語からロシア語へ人手による翻訳を行った。翻訳者は、東京外国語大学大学院博士前期課程の、翻訳家を志望するロシア語母語話者の学生（第2著者）である。

翻訳にあたり、日本語とロシア語の言語構造の違い等により、翻訳に困難を生じさせるであろう箇所が多くあることが予想されたため、一定の方針を設定した。

まず、談話レベルではロシア語としての自然さは失われてもよいとしたが、単文レベルでは自然なロシア語になるような翻訳を行った。例えば、日本語では現在形と過去形が混ざっている文章が多々あるが、ロシア語だと特定の文脈がない限りどちらかで統一するのが一般的である。今回の翻訳では、単文レベルで翻訳元の日本語の文の時制を、ロシア語文でも用いることとした。例えば、(1)のような文を見られたい。翻訳元の日本語文(1a)では、現在形(ル形)と過去形(タ形)が共に用いられており、ロシア語文(1b)でもそれに合わせて翻訳されている。

- (1) a. [...] 二、三年時に担任だった池田弘子先生（七十五）は違った。「そんな薄いかばんじゃ遊び道具も入らないよ」「体育や部活では、危ないからピアスをはずしたほうがいい」。やんわり語りかける。

表3 ロシア語翻訳データと他の言語の翻訳データの比較

	総語数	対象のサンプル数	対象のサンプル名
ロシア語	13,070 語	16 サンプル	PN1c_00001 / PN1d_00001 / PN1e_00001 / PN1f_00001 / PN2e_00001 / PN3b_00001 / PN3g_00001 / PN4a_00001 / PN4b_00001 / PN4c_00001 / PN4f_00001 / PN4g_00001 / PN1a_00002 / PN1b_00002 / PN1d_00002 / PN1e_00002
英語	4,840 語	6 サンプル	OY04_00001 / OC01_00001 / PM25_00001 / PB12_00001 / PN1c_00001 / OW6X_00000
イタリア語	6,563 語	16 サンプル	OC01_00001 / OW6X_00000 / OY04_00001 / PB12_00001 / PM25_00001 / PN1c_00001 / OC02_00001 / OY12_00005 / OC03_00001 / OY09_00008 / OC04_00001 / OY15_00014 / OC05_00001 / OY04_00017 / OC06_00001 / OY04_00027
インドネシア語	51 語	1 サンプル	OY09_00008
中国語	7,852 字	6 サンプル	OY04_00001 / OC01_00001 / PM25_00001 / PB12_00001 / PN1c_00001 / OW6X_00000

- b. Xiroko Ikéda (75 let), kotoraja byla ee klassnym [Hiroko Ikeda]-NOM.F 75 ages who-NOM.F be-PST.F her [homeroom rukovoditelem na vtorom i tret'em godu obučenija, teacher]-INS on second-LOC and third-LOC year-LOC education-GEN byla ne takoj. Ona razgovarivaet s devočkoj mjangko: <<V be-PST.F not such she-NOM talk-PRS.SG.3 with girl-INS softly in takuju ploskuju sumku daže igry ne vlezut>>, <<Na such-ACC thin-ACC bag-ACC even toys-NOM.PL not go-PRS.PL.3 on fizkul'ture i v sekcijax lučše snimat' serežki, physical\_education-LOC and in section-LOC better take\_off-INF earrings-ACC

éto opasno}}.

it\_is dangerous <sup>(3)</sup>

(読売新聞 [ BCCWJ: PN1c.00001 ])

また、ロシア語では、ロシア国外の企業名や国外の新聞の名前などはキリル文字で表記される場合とローマ字で表記される場合がある。今回の翻訳では、企業名等は特定の場合<sup>(4)</sup>を除き、ローマ字で表記することとした。なお、このキリル文字とローマ字の表記については、単に表記の問題というわけではなく、文法面にも影響を与える問題である。一般に、ローマ字表記となる場合、その名詞は曲用しないことになるが、キリル文字表記であれば曲用する<sup>(5)</sup>。ロシア語では固有名詞でもキリル文字表記であれば曲用してしまうため、ローマ字表記にして曲用させない方がもともとの名称がわかりやすい。

固有名詞には、上位概念を示す名詞 (例えば、*kompanija* 「会社」、*proizvoditel'* 「メーカー」など) が同格句として前置される場合もよくある<sup>(6)</sup>。今回の翻訳では、(2a) のように上位概念を示す名詞が日本語で表示されていても、(3a) のように表示されていなくても、(2b, 3b) で示すようにロシア語ではこのような上位概念を示す名詞を加えることとした。

(2) a. 米ガートナー・グループ傘下の調査会社データクエスト

b. kompanija Dataquest, prinadležaščaja amerikanskoj Gartner  
company-NOM Dataquest belong\_to-PTCP.NOM American-DAT Gartner  
Group  
Group

(産経新聞 [ BCCWJ: PN1d.00002 ])

(3) a. ファミリーマートは9日、[...] 発表した。

b. 9 čisla kompanija FamilyMart soobščila [...].  
9 number company-NOM.F FamilyMart announced-PST.F

(産経新聞 [ BCCWJ: PN1d.00002 ])

(3) キリル文字はローマ字に翻字する。翻字は以下の通りである：A=A, Б= B, В=V, Г=G, Д=D, Е=E, Ё=E, Ж=Ž, З=Z, И=I, Ы=J, К=K, Л=L, М=M, Н=N, О=O, П=P, Р=R, С=S, Т=T, У=U, Ф=F, Х=X, Ц=C, Ч=Č, Ш=Š, Щ=Šč, Ь=”, Ы=Y, Ь=’, Э=É, Ю=Ju, Я=Ja. また、本稿で用いる文法情報の略記は以下の通りである：NOM=主格, GEN=属格 (生格), DAT=与格, ACC=対格, INS=具格 (造格), LOC=前置格 (処格), M=男性, F=女性, N=中性, PRS=現在, PST=過去, PTCP=分詞 (形動詞), INF=不定形, 3=3 人称。

(4) 例外となる特定の場合とは、ロシア語で正式名称のあるものである。例えば、国外の新聞の名前で言えば、「人民日報」(*Žén'min' žibao*)、「ルモンド」(*Mond*)、「タイムズ」(*Tajms*)、「エスタド・デ・サンパウロ」(*Éštadau*) などである。

(5) 例えば、*Toyota* 「トヨタ」であれば、ローマ字表記の場合は、格により形態が変化することはないが、キリル文字で表記した場合、(i) で示すように曲用する。

(i) Tojota, Tojoty, Tojote, Tojotu, Tojotoj, o Tojote  
Toyota-NOM Toyota-GEN Toyota-DAT Toyota-ACC Toyota-INS about Toyota-LOC  
「トヨタが, トヨタの, トヨタへ, トヨタを, トヨタによって, トヨタについて」

(6) 以降、(2-6) では、上位概念を示す名詞に下線を引く。

(3b) では、上位概念を示す名詞として *kompanija* 「会社」が付加されている。もし、これが示されていない場合、*FamilyMart* を知っているロシア語母語話者は一つの店舗としての *FamilyMart* を想定してしまう可能性が高い。この場合は、(4) のように上位概念を示す名詞として *magazin* 「店」を想定することになる。*magazin* が表示されていなくても、動詞は形式上これと一致し、男性形となる<sup>(7)</sup>。

- (4) 9 čisla (magazin) FamilyMart soobščil [...].  
9 number shop-NOM.M FamilyMart announced-PST.M

このように、上位概念を示す名詞の違いによって、内容的な齟齬をきたすだけでなく、文法面にも影響が出る。適切な上位概念を示す名詞を付加することで曖昧性を排除することができ、文がわかりやすくなるという効果がある。

地名については事情が少々複雑である。(5a) で示すように、上位概念を示す名詞があっても、ロシアの地名は普通は曲用する<sup>(8)</sup>。これと同様の方針で翻訳すれば、例えば「静岡県」は(5b) のように曲用させることになる。

- (5) a. v gorode Moskve in city-LOC Moscow-LOC 「モスクワ(という都市)で」  
b. v prefektуре Sidzuoke in prefecture-LOC Shizuoka-LOC 「静岡県で」

しかしながら、この場合では、もともとの名称が分かりにくい。よって、今回の翻訳では、上位概念を示す名詞を付加し、(6) のように地名そのものは曲用させない方針とした。

- (6) a. 静岡県出身。  
b. Rodom iz prefektury Sidzuoka.  
birth-INS from prefecture-GEN Shizuoka-NOM  
(西日本新聞 [ BCCWJ: PN3g\_00001 ])

ただし、今回の翻訳の対象は PN コアデータであり、出典が新聞からのテキストであることから、地名が多く出現する。地名が出るたび、そのすべてに上位概念を示す名詞 (*gorod* 「都市」、*prefektura* 「県」など) を付加していくのは明らかに文章として不自然となる。よって、「東京・大阪・京都・広島」など、ロシア人にとってもなじみのある地名には(7) で示すように、基本的に日本語文に示されていない限り、このような語は付加しないこととした。

<sup>(7)</sup> ただし、このあたりのロシア語母語話者の言語感覚やロシア語の言語現実は大変複雑であるため、詳細はここでは述べない。

<sup>(8)</sup> ただし、以下(ii) で示すように、河川の名称など、曲用しない場合もあり得る。

- (ii) a. na reke Volge at river-LOC.F Volga-LOC.F 「ヴォルガ川で」  
b. na reke Enisej at river-LOC.F Yenisei-NOM.M 「エニセイ川で」

上位概念を示す名詞(ここでは *reka* 「川」) と名称を示す名詞の性が一致している場合、曲用し(ii a)、一致しない場合は曲用しない(ii b) とされるが、詳細はここでは述べない。

- (7) a. 東京のヨドバシカメラ新宿西口本店  
 b. v glavnom magazine Yodobashi Camera v Tokio u zapadnogo vyxoda  
 in main-LOC shop-LOC Yodobashi Camera in Tokyo at western-GEN exit-GEN  
 so stancii Sindzjuku  
 from station-GEN Shinjuku  
 (朝日新聞 [ BCCWJ: PN4a.00001 ])

この場合, (7) の *Tokio* 「東京」や *Kioto* 「京都」など *-o* で終わる地名は不変化名詞 (indeclinable noun) となるため, 曲用させない。しかし, *Osaka* 「大阪」や *Xirosima* 「広島」のように *-a* で終わる地名は (8) で示すように曲用させることになる。

- (8) a. 広島、大阪各高裁長官を経て  
 b. zanimal post glavy Vysšego suda Xirosimy,  
 was\_engaged\_in post-ACC head-GEN high-GEN court-GEN Hiroshima-GEN  
 zatem glavy Vysšego suda Osaki  
 then head-GEN high-GEN court-GEN Osaka-GEN  
 (西日本新聞 [ BCCWJ: PN3g.00001 ])

#### 4. 日露対照研究への活用の可能性

BCCWJ のロシア語翻訳データの構築により, 日本語の原データと並べることで疑似的な日露対訳コーパスとしての利用も可能であり, 本データは日露対照研究へ活用できると考えられる。本稿では, その一例として, 日本語の文末形式について, 簡単にロシア語と対照させて論じる。

(9a-11a) の日本語の各文の文末形式と (9b-11b) のロシア語におけるその対応部分 (下線部) を見られたい。

- (9) a. [...] 非常通報装置が作動。 [...] 商品のビデオカメラ六十台とノートパソコン四台 [...] が盗まれていた。  
 b. [...] srabotala sistema signalizacii. [...] ukradeno 60  
 worked-PST.F system-NOM.F signaling-GEN stolen-PTCP 60  
 videokamer i 4 noutbuka iz čisla tovarov [...]  
 video\_camera and 4 laptop from number products-GEN  
 (中日新聞 [ BCCWJ: PN4f.00001 ])

- (10) a. [...] 異国の食文化をどん欲に吸収。 [...] パスタもレパートリーに加えた。 [...] 心を奪われた。 [...] 日本語学校にも通い始めた。  
 b. [...] on [...] žadno vpityval kulinarne tradicii  
 he-NOM.M greedily absorb-PST.M culinary-ACC traditions-ACC  
 čužoj strany. On takže dobavil v svoj  
 foreign-GEN country-GEN he-NOM.M also added-PST.M into self's-ACC

kulinarnyj repertuar pastu [...] on byl  
 culinary-ACC repertory-ACC pasta-ACC he-NOM.M was-PST.M  
 neverojatno očarovan [...] on daže načal xodit' v  
 unbelievably fascinated-PTCP.M he-NOM.M even start-PST.M go-INF to  
 školu japonskogo jazyka.  
 school-ACC Japanese-GEN language-GEN

(読売新聞 [ BCCWJ: PN4c.00001 ])

- (11) a. 日本政府は [...] 無形文化遺産保護条約を締結した。 [...] 佐藤禎一大使が [...] 締約  
 受諾書を提出した。同条約は [...] 採択された。締結はアルジェリアなどに続いて  
 三カ国目。

b. [...] pravitel'stvo Japonii prinjalo Konvenciju ob  
 government-NOM.N Japan-GEN accepted-PST.N convention-ACC about  
 oxrane nematerial'nogo kul'turnogo nasledija, [...] protection-LOC intangible-GEN cultural-GEN heritage-GEN  
 Posol [...] Téjiti Sato pred"javil [...] ambassador-NOM.M [Teiichi Sato]-NOM.M presented-PST.M  
 dokument o soglasii na prinjatie konvencii.  
 document-ACC about agreement-LOC on acceptance-ACC convention-GEN  
 Éta konvencija byla utverždena [...] Vsled za this-NOM.F convention-NOM.F was-PST.F approved-PTCP.F following  
 Alžiom i t.d. Japonija stala tret'im gosudarstvom, Algeria and so on Japan-NOM.F became-PST.F third-INS nation-INS  
 prinjavšim konvenciju.  
 taking-PTCP.INS convention-ACC

(西日本新聞 [ BCCWJ: PN4g.00001 ])

(9a-11a) の日本語文の文末形式に注目すると、名詞で文を終える体言止めを用いている箇所がある。一方、日本語文で体言止めとなっている箇所の対応部分では (9b-11b) のロシア語では動詞の過去形 (-l/-la) で示されている<sup>(9)</sup>。

もし、日本語文で体言止めとなっている箇所を完全な文の形式とするのであれば、動名詞 (verbal noun) で体言止めにされている (9a, 10a) では「-した」を追加し、(11a) では例えば「-であった」とコピュラを追加することになる。このようにした場合、その文末形式は、例えば (10a) では「[...] 吸収した。 [...] 加えた。 [...] 奪われた。 [...] 始めた。」となり、過去形 (タ形) のみが続き文章が単調になってしまう。これを避けるために、適宜体言止めが用いられているものと考えられる<sup>(10)</sup>。ロシア語にも、日本語の動名詞を用いた構文のように体系的に動

<sup>(9)</sup> ボールド体で示してある。

<sup>(10)</sup> 出典のテキストが新聞からであるため、字数の制限等も関係してくる可能性がある。また、これについて結論を出すためにはより詳細な検討が必要となる。

詞を名詞化させる方法は存在するが、(9b-11b)で示すように、そのような形式はここでは用いられていない。(9a-11a)の日本語文の文末形式に対応する部分は(9b-11b)のロシア語では1例を除き全て動詞の過去形(-l/-la/-lo)となっている<sup>(11)</sup>が、ロシア語の基本語順はSVOであり(Isačenko 1966など)、基本的に述語動詞は文末に位置しない<sup>(12)</sup>ため、動詞の時制により文末形式が固定されることはない。そのため、文末形式を多様化するという日本語のような理由では、動詞の名詞化の構文は用いられないと考えられる。

以上のように、ロシア語では、文末に述語が位置しないことが多く、動詞の時制が文章内で同一であっても文末形式が固定されることはない。一方、日本語では述語がほぼ必ず文末に位置するために、動詞の時制が過去で統一されてしまうと、文末形式が「タ」に固定されてしまう。日本語では過去形(タ形)の連続を避け、文末形式を多様化するために、体言止めが適宜用いられるといえる<sup>(13)</sup>。

## 5. おわりに: まとめと今後の課題

本稿では、BCCWJのロシア語翻訳データの構築について述べた。

BCCWJの新聞(PN)コアデータ16サンプルを対象に日本語からロシア語へ人手による翻訳を行った結果、ロシア語翻訳データの総語数は13,070語となり、既に構築されていた英語、イタリア語、インドネシア語、中国語の各翻訳データと比べると最大の規模となった。

翻訳の際は、日本語文の時制をロシア語文でも用いること、企業名等はローマ字で表記すること、固有名詞にはできる限り上位概念を示す名詞を付加すること、「東京・大阪・京都・広島」などのロシア人もよく知る地名に限り上位概念を示す名詞を付加しないこと、等を方針とした。

コーパス研究が盛んである今日でも日本語・ロシア語の対訳コーパスは大変希少であり、BCCWJの16サンプルであっても、その基礎となり得るデータを構築したことは日露対照研究、さらには類型論研究に対し一定の意義のある言語資料を提供できたといえる。本稿では、翻訳データの日露対照研究への活用の一例として、日本語の文末形式について、簡単にロシア語と対照させて論じた。ロシア語とは異なり、日本語の新聞には体言止めがしばしば使用されることを指摘し、その要因は文末形式を多様化するためであるとした。

Soejima (2017)は文学作品(とその翻訳作品)を用いて日露語の対訳コーパスを構築し、不特定の動作主が関わる意図的な出来事が日本語とロシア語でどのように表現されるかについて検討している。その結果、過程を表す場合は、日本語では受動文がよく用いられる一方でロシア語では不定人称文がよく用いられるとしている。結果を表す場合は、日本語では自動詞文がよく用いられ、ロシア語では受動文や自動詞文など多様な形式が用いられるとしている。Soejima (2017)は文学作品でこの結論を導いているため、新聞という異なるレジスターでも同様の結果が得られるか今後検討したい。

さらに、今回ロシア語に翻訳した日本語のサンプルは情報構造のアノテーションもなされて

(11) (9b)の *ukradeno*「盗まれた」のみ受動分詞である。

(12) ただし、ロシア語は語順が自由であり、述語動詞を文末に位置させることも可能ではある。

(13) もちろん、これについて結論を出すためにはより詳細な検討が必要となる。

いる (Miyouchi et al. 2018). よって, 今回構築したロシア語翻訳データにも情報構造のアノテーションを施せば, 日本語・ロシア語の情報構造についての対照研究が量的に行えるようになる<sup>(14)</sup>. 日本語もロシア語も共に顕在的な冠詞のない言語であるため, 定性や特定性などの冠詞を持つ言語では冠詞が担う機能をどのように表現するか (またはしないか)<sup>(15)</sup>を今後詳細に検討したい.

## 謝 辞

本研究は国立国語研究所コーパス開発センターの共同研究プロジェクト「コーパスアノテーションの拡張・統合・自動化に関する基礎研究」(プロジェクトリーダー: 浅原正幸)の研究成果である. また, JSPS 科研費 (課題番号: 17J07534) の助成を受けている.

## 文 献

- Kikuo Maekawa, Makoto Yamazaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka, and Yasuharu Den (2014). “Balanced corpus of contemporary written Japanese.” *Language Resources and Evaluation*, 48:2, pp. 345–371.
- Aleksandr V. Isačenko (1966). “O grammatičeskom porjadke slov.” *Voprosy Jazykoznanija*, 6, pp. 27–34.
- Kensaku Soejima (2017). “On expressions of agent de-topicalized intentional events: A contrastive study between Japanese and Russian.” *Journal of Japanese Linguistics*, 30:1, pp. 107–128.
- Takuya Miyouchi, Masayuki Asahara, Natsuko Nakagawa, and Sachi Kato (2018). “Information-Structure Annotation of the “Balanced Corpus of Contemporary Written Japanese”.” *Computational Linguistics Vol. 781. Communications in Computer and Information Science*, pp. 155–165. Singapore: Springer.
- Masayuki Asahara (2017). “Between Reading Time and Information Structure.” *Proceedings of The 31st Pacific Asia Conference on Language, Information and Computation PACLIC 31*.
- Catherine V. Chvany (1973). “Notes on ‘root’ and ‘structure-preserving’ in Russian.” C. Corum, T.C. Smith-Stark, and A. Weiser (Eds.), *You take the high node and I will take the low node*. Chicago, IL: Chicago Linguistic Society. pp. 252–290.
- Alan Timberlake (1975). “Hierarchies in the Genitive of Negation.” *The Slavic and East European Journal*, 19:2, pp. 123–138.

<sup>(14)</sup> さらに, Asahara (2017) は, 情報構造と読み時間の関係を考察している. ロシア語のデータに対するアノテーションを充実させれば, 同様の研究が可能となり, Asahara (2017) で示された成果が日本語特有のものか通言語的なものかを調べることも可能となる.

<sup>(15)</sup> 一般にロシア語は定性を語順によって表現する傾向があるとされる (Chvany 1973 など) が, あくまで傾向に過ぎない. また, 否定属格の現象も定性との関係がよく指摘される (Timberlake 1975 など) が, これもはっきりとした規則ではない.