

狂言対照コーパス構築における問題点と対策

北崎 勇帆（東京大学大学院 人文社会系研究科）

同一のあらすじ・系統を持ちながら、成立年代の異なる資料組の差分を精査することによって言語の時代差を見るという手法が、言語史研究においてしばしば取られる。本稿では、そのような研究に資するための対照コーパスを構築することを目的とし、アライメントを行う際に発生し得る問題と対策について述べる。その際、狂言台本の組を具体的な事例とし、言語史研究のための対照コーパス構築一般に起こり得る問題と、狂言台本に特有の問題に切り分けて論じる。

Constructing a Parallel Corpus for *Kyōgen* Scripts: Issues and Solutions

Kitazaki Yuho (Faculty of Humanities, the University of Tokyo)

In the field of historical linguistics, an approach often used involves comparing a set of the articles having a similar outline but written in different eras. This paper describes how to design and construct a parallel corpus to contribute to this kind of study. As an example of such articles, this paper takes up *kyōgen* scripts. These scripts contain problems generally attributed to a parallel corpus and peculiar to *kyōgen* scripts.

1. はじめに

日本語史研究において、例えばある形式の使われ方の変遷を調べる場合に、同一のあらすじを持ちながら（成立年代が異なるなどの理由によって）異なる本文を持つ資料組の、対応する箇所を比較するという手法が取られることがある。例えば、『天草版平家物語』（1592年刊）とその原拠本とされる『覚一本平家物語』『斯道文庫本百二十句本平家物語』（いずれも鎌倉頃成立）の比較、朝鮮で編纂された日本語学習用のテキスト『捷解新語』（1636頃成立・1676刊）とその改修本（1748刊）、重刊改修本（1781刊）の比較などがその事例である。また、年代の異なりに着目するほか、文体の異なりに注目して『今昔物語集』と『宇治拾遺物語』に共通する説話を対象として文体差を抜き出す[1]、江戸・上方間で改作が行われた近世洒落本を対象として東西方言の差を炙り出す[2]といった事例もある。このような「同一のあらすじ」かつ「異なる本文」を持つテキスト組を資料とする研究は「何が何に置き換えられたか」「何と何が対応関係にあるか」という点が明確であるために、時代差・文体差・位相差を求めやすいというメリットを持つが、目視によって対応箇所を一つ一つ確認しなければならないこと、調査範囲が広範に亘ることなどから、概して特定の形式の相違に注目した研究に偏りがちであり、総体的な検討を行うことは難しかった。

こうした背景から、稿者は北崎（2015a, 2015b）[3,4]において、同一のあらすじ且つ異なる本文を持つ資料組を対象とした対照コーパスの設計についての提案を行った。具体的には、北崎（2015a）では、TEI P5 Guideline[5]に基づいて資料をXMLデータ化し、文ごとにIDを付与して、対応箇所を属性値として与えることによって対照コーパス構築を実現できることを論じた。ここでは対応

付けの自動化の手法については未検討であったため、北崎（2015b）では、特に狂言台本組を事例として、語ごとのアライメントを動的計画法によって行うことを提案し、特に、表層形ではなく語彙素をベースとして比較することで対応付けの精度が向上することを述べた。

北崎（2015b）で対象とした大蔵流狂言台本2本のうち、『虎明本』（とらあきらぼん、全236曲、1642年大蔵虎明書写、室町期の言語を反映するとされる）は、形態論情報が付与されたタグ付きコーパスが『日本語歴史コーパス室町時代編Ⅰ狂言』として公開されている[6]が、『虎寛本』（とらひろぼん、全165曲、1792年大蔵虎寛書写、現行狂言に近く、近世前期の言語現象を反映するとされる）はこれまでテキストデータ化などの電子化がされたことのない資料であった。そこで稿者は今夏、笹野堅校訂（1942-1945）[7]を底本とした電子テキスト（約60万字）を完成させた。理想的には、これらのテキスト中の文・語が、それぞれ対応付けられることが望ましく、そのような対照コーパスが実現された場合、例えばある語形・形式やある意味範疇を示す形式の一群が、室町時代から近世前期にかけてどのように変化したか、という問題の解決に寄与するであろう。

しかしながら、実際に当該資料組に手法を適用しようとする、対照コーパス構築における汎用的な問題、言語史研究に資するためのコーパス特有の問題、狂言台本特有の問題など、様々な問題が発生した。本稿では、こうした問題を報告しつつ、併せて、現段階で考えられる対応策について述べたい。

以降、第2節で、本稿で用いる底本と使用データの概説を行い、第3節で文対応の問題について、第4節で語対応の問題について論じ、第5節で展望と今後の課題を述べる。

2. 使用データと設計

まず、今回使用する『虎明本』『虎寛本』について説明する。『虎明本』は台本の形式を持つ中では最古の狂言台本、大蔵虎明の1642年の書写によるもので、全236番を収録する。大塚光信編(2006)[8]を底本として、形態論情報が付与されたタグ付きコーパスが『日本語歴史コーパス 室町時代編 1 狂言』として、国立国語研究所より公開されており、今回はそのデータを利用した。『虎寛本』は『虎明本』から150年下の1792年に大蔵虎寛によって書写された台本であり、大蔵流の現行狂言とほぼ同じ本文を持つとされる。『虎寛本』所収の曲は概ね『虎明本』にも収録されており、所収曲は全165番と『虎明本』に比して少なく、整理された状況を窺わせる。書写年代は江戸後期であるものの、言語現象は近世前期のものを反映しているとされる[9]。この2本の狂言テキストを対照させることによって中世・近世間の言語変遷を観察する試みは多く、これまでに一定の成果が上がっている[10,11,12]。『虎寛本』を翻刻したものとして、笹野堅校訂『能狂言 大蔵虎寛本』(岩波書店)があり、今般その電子テキスト化を行ったことは、前節にも述べた通りである。

サンプルとして、『虎明本』『虎寛本』それぞれの「夷大黒」[a]の冒頭部を示す。以下、断りのない場合は引用は全て「夷大黒」による。

・虎明本「夷大黒」冒頭部

「次第おとこ、又なのりばかりにてでもする、又次第にてすれば、やどへ帰りつゐたと云て、くわんじやうする、なのり計にては左のごとく、『かへるうれしき古郷に、／＼、いそぎて、めこにあわうよ』『是は津国あしやの里の者にて候、あまりに浮世を送りかね、ひえの山の三面の大黒は、いづれの大こくよりもれいげんあらたにて、いのりをかけ申せば、程なく富貴に守り給ふと申程に、ひえいざんへのぼり、一夏を送りて候へば、御夢想に、にしの宮へ参て、ゑびす三郎殿へ祈誓をして、吉日をもつて、くわんじやうせよとの御夢想にて候間、西の宮へ参り祈念を申てあれば、是も吉日を以て勸請せよとの御つげにて候、かやうにありがたき事はあるまじく候、則今日吉日にて候、是へくわんじやう申、みしめをひかばやとぞんずる

・虎寛本「夷大黒」冒頭部

(アド、謡) [次第] 帰る嬉しき古郷に、／＼、いそいで妻子に逢ふよ。(アド) 是は津の国蘆屋の里の者で御座る。某比叡山三面の大黒天と、西

の宮の夷三郎殿へ、楽しく成して被下いと祈誓を掛て御されば、吉日をえらび、勸請せよとの御示験で御さる。すなはち今日は吉日で御さるに依て、注連をひき、勸請致うと存る。

構造化にあたっては、まず、文・語にそれぞれIDを付与し、さらに、それぞれの文・語についてどの箇所に対応するか(もしくはしないか)をタグ付けすることとする(田中 2013[13]も参照)。ここでは『虎明本』側をベースとして、対照用の情報を付与していく。

3. 文対応の問題

3. 1 文対応の問題 (1)

北崎(2015b)は文対応が取れていることを前提とするため、まずは文アライメントを行う必要がある。例えば、次のような文対応の関係が取れていることが、本来的には望ましい。図1にはそれぞれの「夷大黒」の一部を対照し、さらに、簡易的なタグ付けを行ったものを示す[b]。

ID	虎明本	ID	虎寛本
23	をとこ	20	(アド)
24	「是はかたじけな ひ御事にて候、	21	ハア、是は有難う 御座る。
25	又それにみえさせ 給ふは、いかやう なる御かたにて候 ぞ	22	又あれにたゝせら れたはいか様な御 方で御座る。

↓

```
<s id="23" corresp="20">をとこ</s>
<s id="24" corresp="21">「是はかたじけなひ御事にて候、</s>
<s id="25" corresp="22">又それにみえさせ給ふは、…</s>
```

図1 理想的な文対応とXML化

しかし、以下に示すように、狂言台本組の文は必ずしも1対1で対応しないか、または、そもそも対応する本文がもう片方の台本に存在しないこともある。

ID	虎明本	ID	虎寛本
10	則今日吉日にて 候、	5	すなはち今日は吉 日で御さるに依
11	是へくわんじやう 申、みしめをひか ばやとぞんずる		て、注連をひき、 勸請致うと存る。

図2 1対多対応の場合

^a ある家の主人が比叡山の大黒天と西宮の夷三郎の来臨を願うと、両神が現れてそれぞれ由来を語る。大黒・夷ともに主人に宝物を与え、めでたく舞う、というあらすじである。

^b `<s>`は文の単位を、属性値`id`は文のID(通し番号)を、属性値`corresp`は対応する『虎寛本』の文のIDを示す。`corresp`はTEIガイドラインでは“(corresponds) points to elements that correspond to the current element in some way.”と定義される。

ID	虎明本	ID	虎寛本
20	ゑびす	14	(夷)
21	「いかやうなる 者ぞとはおろか なるとひ事かな,	15	某を得しらぬか.
	(対応なし)	16	(アド)
		17	何共存ませぬ.
		18	(夷)
22	くわんじやうせ よとちげんをお ろしたる, 西の宮 のゑびす三郎殿 にてあるぞとよ	19	是は西の宮の夷 三郎殿なるが,現れ出て有る ぞとよ.

図3 対応箇所がない場合

図2には『虎明本』の2文が『虎寛本』の1文に相当する場合, 図3には『虎寛本』におけるアド^aの返答が『虎明本』に見られない場合を例として挙げた. このような場合には1対1の文対応を前提とすることはできない.

なお, これらの問題は, 典型的には要約文と元テキストの文対応[14,15]にも同様に見られる問題であるため, アライメントの際には同様の手法を適用することによって解決できることが予想される.

3. 2 文対応の問題 (2)

さらに, おそらくは狂言台本に特有のものと見られるものに, 「台本中の繰り返し箇所の参照」の問題がある. 図4において, 『虎明本』の「夷大黒」におけるID44から55までの箇所は, 『虎寛本』においてはID61の箇所で「語りの箇所は「大黒連歌」と同様」である旨が注記され, 省略が行われている. すなわち, 「夷大黒」の本文同士を対照するだけでは, この箇所の比較を行うことはできず, 他の曲(ここでは「大黒連歌」)の当該箇所を参照しなければならない. 以下に, 虎明本の「夷大黒」における大黒の由来を語る場面と, 虎寛本の「大黒連歌」における当該箇所を挙げる.

・虎明本「夷大黒」の大黒の由来を語る場面

かたり「そも / \ ひえいざんゑんりやくくじは, 伝教大師くわんむ天王と御心をひとつにして, ゑんりやく年中にかいひやくし給ふ, されば一ねん三ぜんの機をもつて, 三千人の衆徒を置, 仏法今にはんじやうたり, 其時伝教大師此山には, 三千の衆徒あれば, 一日に三千を守り給ふ, てんぶをとときせし給ふ所に, 此大黒出現する, かいさんいひや大黒は, 一日に千人をこそふちし給へ, 此山には三千人の衆徒あれば, 三千を守り給ふ, てんぶをこそあんじ申べけれとありしかば, 此大こ

く大にいかりをなし, いでさらば三千を守る奇特をみせんとて, たちまち三面六ひと現じ, 今におひて仏法はんじやうに守るなり, なんぼうきどくなる大こくにてあるぞ,

・虎寛本「大黒連歌」の大黒の由来を語る場面

(シテ, 語) 抑比叡山延暦寺は, 伝教大師, 桓武天皇と御心をひとつにして, 延りやく年中に開闢し給ふ. さあるに依て寺号を延暦寺と号す. されば一念三千の機を以, 三千人の衆徒を置, 仏法今に繁昌たり. 其時伝教大師, 加程の山に守護神なくては叶はじとて, 一日に三千人を守り給ふ天部をと祈誓し給ふ所に, 此大黒出現す. 開山, いゝや, 大黒は一日に千人をこそ扶持し給へ. 此山には三千人の衆徒有れば, 大黒天はいかゞと有る. 其時大こく大きに怒て, いでさらば三千人を守る奇特を見せんとて, 忽三面六臂と現れければ, 開山喜悦のおもひをなし, 夫よりひえい山無動寺の三面の大黒天といわれ, 今におゐて仏法はんじやうに守るなり.

ID	虎明本	ID	虎寛本
40	おとこ	53	(アド)
41	「近比ありがた う候,	54	是は有難う御ざ る.
42	扱三面の大こく の御ゆらひはい かに	55	大こく天の御由来 をも承り度う御ざ る.
	(対応なし)	56	語て聞せう.
		57	能う聞.
		58	(アド)
		59	ハア.
		60	(シテ)
43	かたり	61	語り大こく連歌同 断.
44	「そも / \ ひえ いざんゑんりや くじは, ... (12文 で大黒天の由来 を説明する)		(対応なし)
	⋮		
56	心やすくしんが うせよ,	62	心安う信仰せよ.
57	たのしうなさう ずるぞ	63	たのしう成して取 らせうぞ.

図4 繰り返し箇所の参照

これは, 狂言台本が一種の類型やテンプレートを持ち, かつ, 両台本に収録される曲の順が一定しないことに起因するものである. 「夷大黒」は『虎明本』では1番目に配される曲目であるが, 『虎寛本』では11番目に位置するため(表1), 大黒の由来を語る場面は「語り大黒連歌同断」と

^a 「アド」は狂言における脇役. ここでは夷(シテ・主役)に対する, 「男」の役を指す.

して前の「大黒連歌」（10 番目）の該当箇所を参照することによる省略が行われているのである。

表 1 虎明本・虎寛本の曲順

	虎明本	虎寛本
1	夷大黒	すゑひろがり
2	連歌毘沙門	目近
3	福の神	麻生
4	大黒連歌	三本の柱
5	此沙門	隠れ笠
6	餅酒	宝の槌
7	かくすい	よろひ
8	昆布柿	福の神
9	鴈鴻金	連歌毘沙門
10	三人夫	大黒連歌
11	筑紫の奥	えびす大こく
⋮	⋮	⋮

すなわち、正確に対応付けるならば、当該箇所は、実際には図 5 のような対応関係を持つ。このような問題への対応を考える場合、これは単に文アライメントの問題に留まらず、「対応」をどのように考えるか、どのように対照コーパスを構築するか、という問題にも関わる。例えば、当該箇所の「夷大黒」のテキスト内に「大黒連歌」の本文を引用し、テキストを組み替えた上で対応付けを行えば、線条性は担保されるが、資料としての信頼度は低下してしまうだろう。

ここでは、テキストそのものに対しては組み換えを行わず、台本全体を通して曲 ID・文 ID・語 ID を付与することにより、原文の配列を保存しつつも、対応箇所を示したコーパスを設計できるものとする。すなわち、対応箇所の文 ID には曲の ID を含めた状態で記述する必要があるため、図 6 のような XML 記述を行うことになる[a]。

4. 語対応の問題

語対応にも文対応と同様、一対多対応・多対多対応の問題がある（田中 2013）。図 1 の箇所をひとまず語ごとに対照すると図 7 のようになるが、351-356「かたじけなひ御事にて候」と 277,278「有難う御座る」や、361-364「みせさせ給ふは」と 283-287「たゝせられたは」の語がそれぞれどう対応するかは、厳密に定義し難い。

動的計画法によるアライメントは距離の近さ・語の類似度を基準として対応付けを求めるため

に、資料間で改変された語の対応付けが困難であるという問題がある。また、語彙素の距離・類似度に基づくアライメントでは 351「かたじけなひ」と 277「有難う」、372「候」と 293「御座る」の対応関係を求めるのは難しい。これは動的計画法を用いる手法の弱点である（山田 2016[16]も参照）。こうした問題については、同一語に注目するのではなく、IBM Model[17]などの統計的機械翻訳のモデルを利用することで、改善されることが予想される。

5. 展望

以上、本稿では、狂言台本を事例として、対照コーパスを構築する際の問題点について述べた。今回は問題の列挙に留まり、実装と公開の詳細に関しては扱うことができなかったが、今回の問題提起を土台として、今後行っていきたいと思います。

なお、どのようにデータを提供するかといった問題について、岩波文庫版『虎寛本』校訂者の笹野堅は 1961 年に没しているため、著作権保護期間は満了している[b]。一方で『虎明本』のデータは底本が 2006 年刊行の編集物であるから、それぞれの本文をそのまま対照併記して公開するといったことはできない。しかしながら、『日本語歴史コーパス（室町時代編）』[6]では全ての形態素に対して一意の ID が付与されている[18]ため、例えば、『日本語歴史コーパス』上で表示された検索結果の ID を入力することで、対応する（もしくは近似する）『虎寛本』の本文箇所を出力することは十分に可能である。また逆に、『虎明本』の対応する、一意の ID をタグとして付与した『虎寛本』のデータを提供することにより、利用者が『日本語歴史コーパス』上で『虎明本』の該当箇所を簡便に検索することも可能になるであろう[c]。

^a もしくは、別曲と対応がある場合にのみ、「当該曲に対応箇所が存在しない」ことを示す属性値を用意するといった方法が考えられるか。

^b そのため、『虎寛本』についてはプレーンテキストの公開が可能である。2018 年度末を目処に、何らかの形で公開を予定している。

^c 検索システム『中納言』上では、一意の ID を URL 内に組み込むことによって用例を permalink として提示することができる（例：https://chunagon.ninjal.ac.jp/chi/permalink?unit=short&position=40-虎明_1642_01001.5890）。この permalink を利用して、『虎寛本』のデータ側からリンクを行うような連携の仕方も考えられる。

虎明本「夷大黒」	虎寛本「夷大黒」	虎寛本「大黒連歌」
42 扱三面の大こくの御ゆらひはいかに	55 大こく天の御由来をも承り度う御ざる、	85 (一のアド)
	56 語て聞せう。	86 扱大黒天の御由来が、
	57 能う聞。	87 (一のアド・二のアド)
	58 (アド)	88 承り度う御ざる。
	59 ハア。	95 (シテ)
	60 (シテ)	96 汝もようきけ。
	61 語り大こく連歌同断。	97 (二のアド)
43 かたり		98 畏て御座る。
44 「そも / ひえいざんゑんりやくじは、		99 (シテ、語)
伝教大師くわんむ天王と御心をひとつにして、		100 抑比叡山延暦寺は、
ゑんりやく年中にかいひやくし給ふ、		伝教大師、桓武天皇と御心をひとつにして、
45 されば一ねん三ぜんの機をもつて、		延りやく年中に開闢し給ふ。
三千人の衆徒を置、仏法今にはんじやうたり、		101 さあるに依て寺号を延暦寺と号す。
		102 されば一念三千の機を以、
46 其時伝教大師		三千人の衆徒を置、仏法今に繁昌たり。
47 此山には、三千の衆徒あれば、		103 其時伝教大師、
一日に三千を守り給ふ、てんぶを		104 加程の山に守護神なくては叶はじ
48 ときせいし給ふ所に、此大黒出現する、		105 とて、一日に三千人を守り給ふ天部を
49 かいさんいひや大こくは、		
50 一日に千人をこそふちし給へ、		106 と祈誓し給ふ所に、此大黒出現す。
51 此山には三千人の衆徒あれば、三千を守り給ふ、		107 開山、いゝや、大黒は
てんぶをこそあんじ申べけれ		108 一日に千人をこそ扶持し給へ。
52 とありしかば、此大こく大にいかりをなし、		109 此山には三千人の衆徒有れば、大黒天はいかゞ
		110 と有る。
53 いでさらば三千を守る奇特をみせん		111 其時大こく大きに怒て、
54 とて、たちまち三面六ひと現じ、		112 いでさらば三千人を守る奇特を見せん
今におひて仏法はんじやうに守るなり、		113 とて、忽三面六臂と現れければ、
		開山喜悦のおもひをなし、
55 なんぼうきどくなる大こくにてあるぞ、		夫よりひえい山無動寺の三面の大黒天といわゝれ、
		今におみて仏法はんじやうに守るなり。
56 心やすくしんがうせよ、	62 心安う信仰せよ、	114 心易う信仰せよ。
57 たのしうなさうずるぞ	63 たのしう成して取らせうぞ、	115 たのしう成して取らせうぞ。

図5 他曲の参照と実際の対応関係

<pre> <s id="001_42" corresp="011_55">扱三面の大こくの御ゆらひはいかに</s> <s id="001_43" corresp="010_99">かたり</s> <s id="001_44" corresp="010_100,101" >「そも / ひえいざんゑんりやくじは、伝教大師くわんむ天王と御心をひとつにして、 ...</s> <s id="001_45" corresp="010_102" >されば一ねん三ぜんの機をもつて、三千人の衆徒を置、仏法今にはんじやうたり、</s> <s id="001_46" corresp="010_103">其時伝教大師</s> <s id="001_47" corresp="010_104,105">此山には、三千の衆徒あれば、一日に三千を守り給ふ、てんぶを</s> <s id="001_48" corresp="010_106">ときせいし給ふ所に、此大黒出現する、</s> <s id="001_49" corresp="010_107">かいさんいひや大こくは、</s> <s id="001_50" corresp="010_108">一日に千人をこそふちし給へ、</s> <s id="001_51" corresp="010_109" >此山には三千人の衆徒あれば、三千を守り給ふ、てんぶをこそあんじ申べけれ</s> <s id="001_52" corresp="010_110,111">とありしかば、此大こく大にいかりをなし、</s> <s id="001_53" corresp="010_112">いでさらば三千を守る奇特をみせん</s> <s id="001_54" corresp="010_113" >とて、たちまち三面六ひと現じ、今におひて仏法はんじやうに守るなり、</s> <s id="001_55">なんぼうきどくなる大こくにてあるぞ、</s> <s id="001_56" corresp="011_62">心やすくしんがうせよ、</s> <s id="001_57" corresp="011_63">たのしうなさうずるぞ</s> </pre>	他曲参照部
--	-------

図6 曲の単位を超えた文の対応付けとそのXML化

文 ID	語 ID	虎明本	文 ID	語 ID	虎寛本
23	347	をところ	20	272	アド
24	348	「 (対応なし)	21	273	ハア
	349	是		274	,
	350	は		275	是
	351	かたじけなひ		276	は
	352	御		277	有難う
	353	事			
	354	に			
	355	て			
	356	候		278	御座る
	357	,		279	.
25	358	又	22	280	又
	359	それ		281	あれ
	360	に		282	に
	361	みえ		283	たゝ
	362	させ		284	せ
	363	給ふ		285	られ
				286	た
	364	は		287	は
	365	,			(対応なし)
	366	いかやう		288	いか様
	367	なる		289	な
	368	御		290	御
	369	かた		291	方
	370	に		292	で
	371	て			
	372	候		293	御座る
	373	ぞ			
				294	.

図7 語単位での対応付け

- 8) 大塚光信編：大蔵虎明能狂言集翻刻注解，清文堂出版（2006）。
- 9) 蜂谷清人：虎明本から虎寛本へ ―語形・用法の変遷とその史的位置についての試論―，狂言台本の国語学的研究，笠間書院，pp.25-52（1977）。
- 10) 柳田征司：虎明本狂言と虎寛本狂言との語彙の比較 ―困惑の気持ちを表わす感情語彙に就いて―，安田女子大学紀要，Vol.1，pp.31-44（1967）。
- 11) 神永正史：虎明本のテアル構文 ―競合という観点から―，筑波日本語研究，Vol.14，pp.35-52（2009）。
- 12) 深津周太：狂言テキストにおける感動詞「シシ申」の歴史的研究，HERSETEC テキスト布置の解釈学的研究と教育，Vol.2，No.2，pp.125-146（2009）。
- 13) 田中牧郎：説話のパラレルコーパスの設計 ―平安・鎌倉時代の文体変異の研究に向けて―，第3回コーパス日本語学ワークショップ予稿集，pp.259-268（2013）。
- 14) H, Jing., K, R, McKeown.: The Decomposition of Human-Written Summary Sentences, *SIGIR '99*, pp.129-136(1999).
- 15) 亀田亮宙・李元・内山清子・武田英明・相澤彰子：論文における要約記述に対応するパラグラフの同定手法，*JSAI2013*，(2013)。
- 16) 山田祐実：日本語史資料を対象とした自動アライメント，奈良先端科学技術大学院大学修士論文（2016）。
- 17) P, F, Brown., et al.: The mathematics of statistical machine translation: Parameter estimation, *Computational linguistics*, 19,2, pp.263-311(1993).
- 18) 国立国語研究所：中納言マニュアル・検索方法，入手先〈http://pj.ninjal.ac.jp/corpus_center/chu-01.html〉（参照 2017-11-10）。

付記

『虎明本』のデータについては，国立国語研究所共同プロジェクト「通時コーパスの設計」の成果を利用して頂きました。また，本発表は科学研究費補助金（特別研究員奨励費 課題番号 16J00119）による成果の一部です。

参考文献

- 1) 田中牧郎・山本啓史：『今昔物語集』と『宇治拾遺物語』の同文説話における語の対応 ―語の文体的価値の記述―，日本語の研究，Vol.10，No.11，pp.16-31（2014）。
- 2) 矢野準：近世後期京坂語に関する一考察，国語学，Vol.107，pp.16-33（1976）。
- 3) 北崎勇帆：洒落本を対象とした東西対照コーパスの設計と構築，人文科学とコンピュータ研究会報告，2015-CH-106（5），pp.1-6（2015）。
- 4) 北崎勇帆：動的計画法を用いた狂言台本の語の対応付け，じんもんこん 2015 論文集，pp.99-104（2015）。
- 5) Text Encoding Initiative：TEI: P5 Guidelines，入手先〈<http://www.tei-c.org/Guidelines/P5/>〉（参照 2017-10-1）。
- 6) 国立国語研究所コーパス開発センター（市村太郎・渡辺由貴ほか）編：日本語歴史コーパス室町時代編Ⅰ 狂言（短単位 1.1 / 長単位データ 1.1，中納言バージョン 2.2.1），入手先〈http://pj.ninjal.ac.jp/corpus_center/chj/muromachi.html〉（参照 2017-10-1）。
- 7) 笹野堅校訂：能狂言 大蔵虎寛本，岩波書店（1942-1945）。