

エラーカテゴリーに基づく  
翻訳学習者の学習過程における習熟度の分析

山本 真佑花（神戸女学院大学大学院 博士前期課程2年）

田辺 希久子（神戸女学院大学）

藤田 篤（情報通信研究機構）

- エラーカテゴリーの紹介
- 決定木の紹介
- 今回の研究内容
- 結果・考察
- まとめ・今後

- エラーカテゴリーの紹介
- 決定木の紹介
- 今回の研究内容
- 結果・考察
- まとめ・今後

# 翻訳エラーのカテゴリー

- MeLLANGE (Castagnoli et al., 2006)
  - 学習者の翻訳における現象の分析
  - 欧州言語のように「近い」言語間の翻訳を想定
- Simplified MeLLANGE (Babych et al., 2012)
  - MeLLANGE のエラー体系を簡略化、一部を修正
  - 翻訳学習者の英日翻訳にも対応 (Toyoshima, 2014) (Toyoshima et al., 2015)
- 上記の体系のさらなる改良版 (Toyoshima et al., forthcoming)
  - 未観察のエラーの除外、一部を細分類
  - 今回の分析に使用

# 今回使用したエラーカテゴリ (Toyoshima et al., forthcoming)

エラーカテゴリ	TD	SD	Term.	Brief/Style
X1 原文内容の欠落	✓	✓		
X2 原文にない情報の付加	✓	✓		
X3 原文内容の歪曲	✓	✓		
X4a 原文表現の押しつけ (未翻訳)	✓	✓		✓
X4b 原文表現の押しつけ (過度な逐語訳)	✓	✓		
X6 曖昧さ未解消	✓	✓		
X7 不適切/一貫しない訳語	✓	✓	✓	✓
X8 不自然なコロケーション	✓			
X9 構文誤り	✓			
X10 前置詞や助詞の誤り	✓			
X11 活用の誤りや数・性などの不一致	✓			
X12 綴り誤り・誤変換	✓			✓
X13 句読法の誤り	✓			✓
X14 レジスタに適さない用語や表現	✓			✓
X15 ぎこちない訳	✓			
X16 結束性違反	✓			

## X1 原文内容の欠落

..., but there are also a lot of possibilities for some more  
**extreme sports**, ...

その他にもたくさん**スポーツ**が楽しめます

→ **extreme**の訳が抜けている

X1

## X2 原文にない情報の付加

The ascent route from here **descends to** the Cosmiques Hut before traversing Mont Blanc du Tacul and ...

ここからの登山ルートは、コスミック小屋へ**下って宿泊した**後、モンブラン・デュ・タキュルと・・・

→ 原文にない「**宿泊する**」という情報が加わっている

X2

### X3 原文内容の歪曲

It is located in the most northern part of the island at the Tanshui River and ...

- × タンショイ川の最も北に位置し、...
- 島の北側でタンショイ川に面し、...

→ 誤訳

X3

## X4a 原文表現の押しつけ(未翻訳)

This way is the most popular but can be prone to danger from **stonefall**.

この道は最も人気があるが、**stonefall** からの危険な傾向もあります。

→ 訳されていない

X4a

## X4b 原文表現の押しつけ(過度な逐語訳)

This way is the most popular but can be prone to danger from stonefall.

この道は最も人気があるが、stonefall からの危険な傾向もあります。

→ 直訳すぎる

X4b

## X6 曖昧さ未解消

A political party is an organization made up of **citizens** ...

政党とは、・・・**国民や市民** によって作られる組織のことです。

→ 2種類の訳語がある

X6

## X7 不適切/一貫しない訳語

... and the best known is the **Tour du Mont Blanc or TMB** ...

一番知られているものが**ツール・ド・モンブランツアー (TMB)**です。

→ **Tour du Mont Blanc or TMB**は  
**ツール・ド・モンブラン(TMB)**と用語集に記載あり

X7

## X8 不自然なコロケーション

A lot of Europe's big **rivers originate** in the Alps, ...

多くのヨーロッパの大きな**河川は**、アルプス山脈から**始まっていて**、...

→ 語と語のつながりがおかしい

X8

## X9 構文誤り

Plans are even more ambitious with hundreds of high rise buildings to be added (including the highest building in the world) and thousands of islands in the form of the world or palms which are for the rich and famous only.

何百もの高層ビル(世界一高い高層ビルを含む)が建設されたり、大富豪や有名人専用の世界地図やヤシの木の形をした何千もの島が作られるなど計画はさらに意欲的です。

→ 「～たり、～たり・・・」にすべき

他: 読点の位置が不適切で修飾関係が不明瞭な場合

X9

## X10 前置詞や助詞の誤り

In summer Interlaken in Switzerland and Innsbruck in Austria are beloved destinations ...

夏はスイスのインターラーケンやオーストリアのインスブルックは、...

→ 助詞がおかしい

X10

## X11 活用の誤りや数・性などの不一致

For most people a visit to the panorama point at the highest station of the cable car does just fine.

ケーブルカーで行ける最も高い駅からの全景が**見れる**ポイントで…

→ 「ら抜き」言葉

他: 「**です・ます**」の統一など

X11

## X12 綴り誤り・誤変換

The city is surrounded by hills and mountains and the area is prone to earthquakes as well.

台北は**岡**や山に囲まれているため地震が起こりやすい地域でもある。

→ 漢字変換ミス(丘)

X12

## X13 句読法の誤り

### <句読点の使い方>

「 」内には句点をつけない。

### <数字>

1桁は全角(1, 2, 3, …)

2桁以上は半角(10, 100, 1000, …)

万以上の単位語は漢字表記(1万2000, 1億2800万)

→ これら指定の句読法に反する場合

X13

## X14 レジスタに適さない用語や表現

### Introduction

#### 序論

→ トラベルガイドなので  
「はじめに」「紹介」などが適切

X14

## X15 ぎこちない訳

the cheapest room costs about a thousand US\$ a night

価格が低い部屋でもアメリカドルで一泊約1000ドル

→ 冗長である

X15

## X16 結束性違反

Dubai is the second-largest of the seven emirates that together constitute the United Arab Emirates. This is one of the fastest ...

ドバイはアラブ首長国連邦を構成する7つの首長国の中で、二番目に大きい首長国です。ドバイは、...

→ 繰り返しのせいで、語彙の結束性がない

X16

- エラーカテゴリーの紹介
- **決定木の紹介**
- 今回の研究内容
- 結果・考察
- まとめ・今後

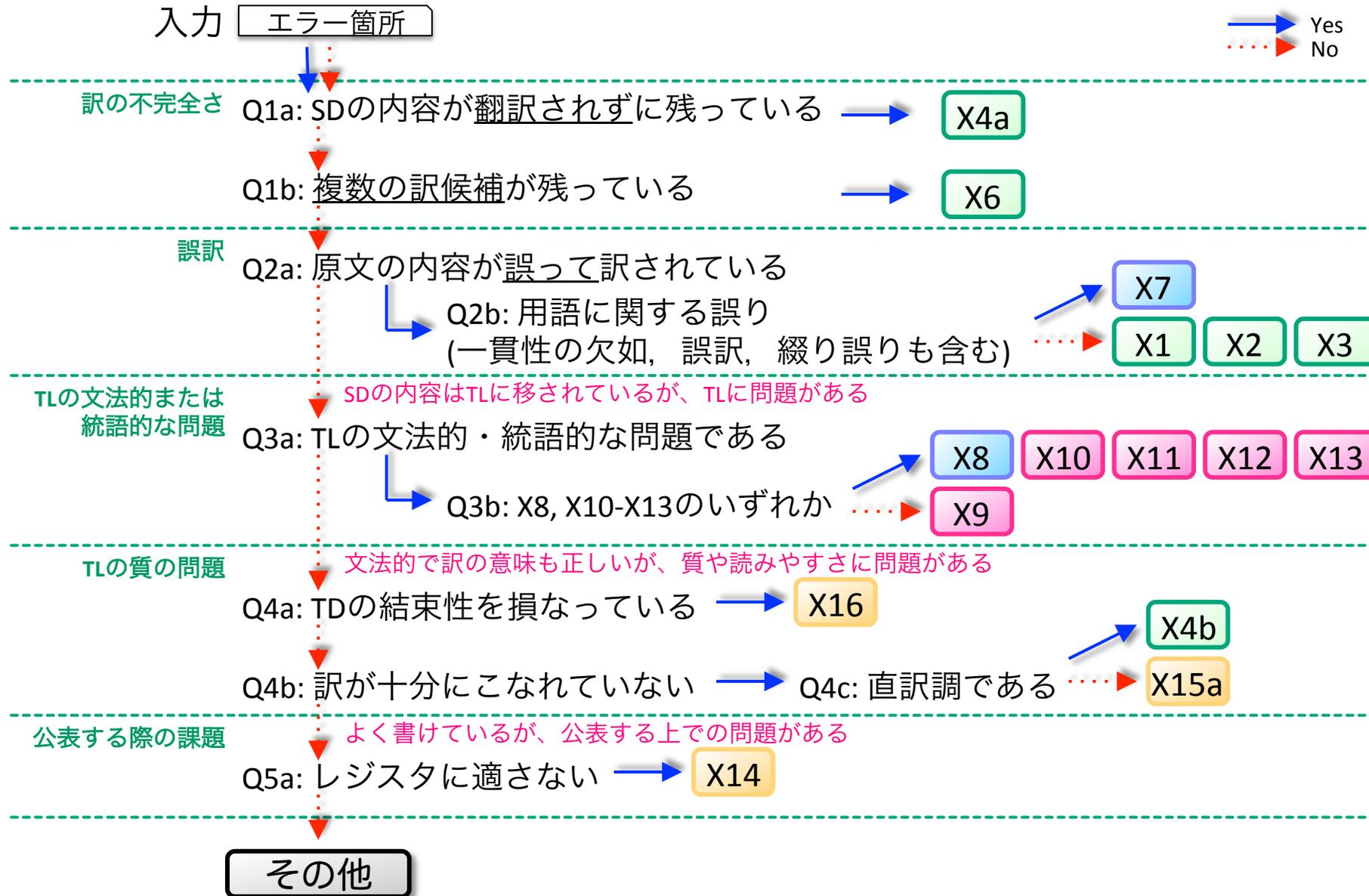
## 決定木作成の経緯

16種類のエラーからひとつを選ぶとなると、  
人によってエラーの解釈にズレが生じる



統一的なエラー付与を可能にするために、  
明確なガイドラインが必要

# エラーカテゴリの決定木 (Toyoshima et al., forthcoming)



- エラーカテゴリーの紹介
- 決定木の紹介
- **今回の研究内容**
- 結果・考察
- まとめ・今後

## 研究目的

習熟過程の中で、

どのエラーが減りやすく、どのエラーが改善困難なまま残るのかを見つけること。



先行研究のように1回のエラー数でなく、半期をかけてエラーの「推移」を見る。

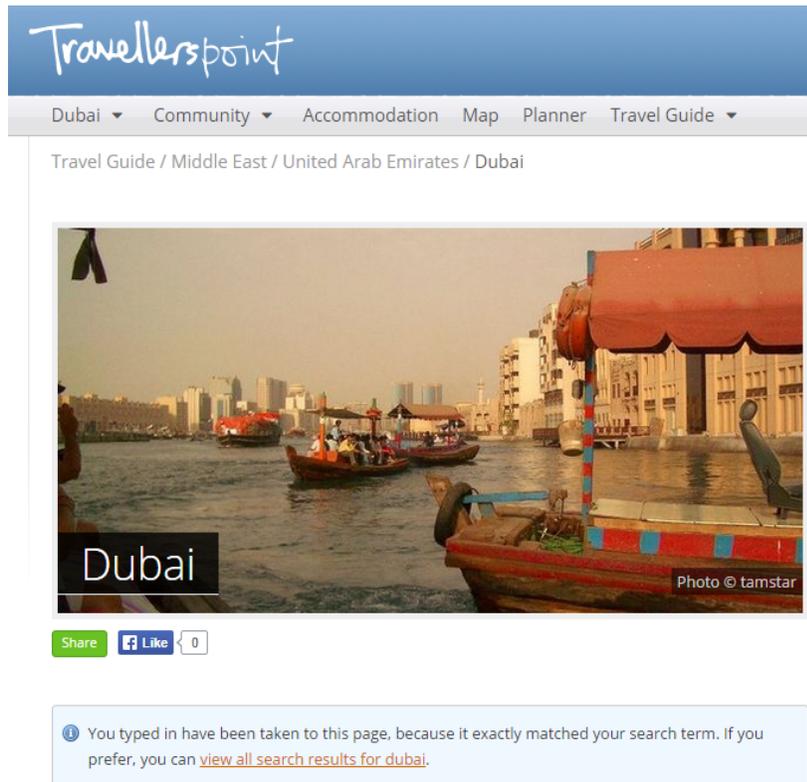
## 今回の研究内容

期間： 2015年4月14日～7月28日

対象学生： A大学在学の大学3、4年生27人

# 使用課題

## Travellerspoint



The screenshot shows the Travellerspoint website interface. At the top, the logo "Travellerspoint" is displayed in a white script font on a blue background. Below the logo is a navigation menu with items: Dubai, Community, Accommodation, Map, Planner, and Travel Guide. The main content area shows a breadcrumb trail: "Travel Guide / Middle East / United Arab Emirates / Dubai". Below this is a large image of a traditional wooden boat (abra) on the water in Dubai, with the city skyline in the background. The word "Dubai" is overlaid on the bottom left of the image, and "Photo © tamstar" is on the bottom right. Below the image are social media sharing buttons for "Share", "Like", and "0". At the bottom, a blue information box contains the text: "You typed in have been taken to this page, because it exactly matched your search term. If you prefer, you can [view all search results for dubai](#)."

## DEMOCRACY NOW!



The screenshot shows the Democracy Now! website interface. At the top, the logo "DEMOCRACY NOW!" is displayed in red and white, featuring a stylized Statue of Liberty. Below the logo is a navigation menu with items: Daily Shows, Top Stories, Web Exclusives, and Topics. The main content area shows the date "THURSDAY, AUGUST 7, 2014" and navigation links "PREVIOUS" and "NEXT". Below this is the article title "Hiroshima and Nagasaki, 69 Years Later". Below the title is a video player showing a woman and a man. Below the video player are download options: "DOWNLOAD: AUDIO FEEDS".

# 各課題の内容

## トラベルガイド

- The Alps  
(192w/195w/184w)
- Dubai  
(180w/210w/197w)
- Taipei  
(168w/167w/209w)

## エイミーズコラム

- Civil Rights  
(155w/203w/151w)
- Free Speech vs. Surveillance in  
the Digital Age  
(178w/155w/205w)
- Hiroshima and Nagasaki  
(175w/216w/164w)

# 研究方法

## 1. 翻訳実習を通じたデータの収集(のべ1725件)

1. 学生によるグロサリー作成、背景知識発表
2. 学生による翻訳
3. 2の結果に対するエラー付与・分類とフィードバック
  - 山本と田辺の2名でダブルチェック
  - 学生ごと、各回ごと、エラーの種類ごとの数を記録

## 2. 分析

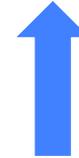
- a. 100ワードあたりのエラー数の推移調査  
(学生ごと、27人全体、エラーカテゴリーごと)
- b. 先行研究([Toyoshima, 2014](#))との誤りの分布の比較
- c. 授業初回と最終回に行ったアンケートの結果との関連付け

- エラーカテゴリーの紹介
- 決定木の紹介
- 今回の研究内容
- **結果・考察**
- まとめ・今後

## a. 100ワードあたりのエラー数の推移調査

Total	Total	4a	6	7	1	2	3	8	10	11	12	13	9	16	4b	15	14
100w_t1	5.19	0.00	0.00	0.95	0.28	0.08	2.50	0.22	0.12	0.04	0.06	0.10	0.36	0.04	0.34	0.06	0.04
100w_t2	5.81	0.02	0.00	1.11	0.56	0.00	1.82	0.06	0.26	0.10	0.21	0.31	0.25	0.11	0.57	0.21	0.22
100w_t3	6.25	0.08	0.00	1.37	0.56	0.19	2.27	0.06	0.13	0.06	0.18	0.49	0.06	0.06	0.66	0.02	0.06
100w_a1	5.88	0.00	0.00	0.63	0.49	0.09	3.33	0.13	0.17	0.05	0.11	0.57	0.06	0.02	0.13	0.00	0.11
100w_a2	5.67	0.02	0.00	0.50	0.52	0.20	2.79	0.04	0.12	0.11	0.11	0.65	0.00	0.07	0.43	0.11	0.00
100w_a3	5.96	0.07	0.00	0.53	0.41	0.23	3.06	0.24	0.22	0.21	0.10	0.47	0.02	0.02	0.16	0.12	0.08

100w_t	5.75	0.03	0.00	1.14	0.47	0.09	2.19	0.11	0.17	0.07	0.15	0.30	0.22	0.07	0.52	0.10	0.11
100w_a	5.77	0.03	0.00	0.54	0.49	0.17	2.99	0.13	0.17	0.12	0.11	0.58	0.03	0.03	0.24	0.08	0.06
Diff	0.02	0.00	0.00	-0.60	0.02	0.08	0.80	0.02	0.00	0.05	-0.04	0.28	-0.19	-0.04	-0.28	-0.02	-0.05



# 減りやすいエラー

X7(一貫しない訳語)、X4b(直訳)

Total	Total	4a	6	7	1	2	3	8	10	11	12	13	9	16	4b	15	14
100w_t1	5.19	0.00	0.00	0.95	0.28	0.08	2.50	0.22	0.12	0.04	0.06	0.10	0.36	0.04	0.34	0.06	0.04
100w_t2	5.81	0.02	0.00	1.11	0.56	0.00	1.82	0.06	0.26	0.10	0.21	0.31	0.25	0.11	0.57	0.21	0.22
100w_t3	6.25	0.08	0.00	1.37	0.56	0.19	2.27	0.06	0.13	0.06	0.18	0.49	0.06	0.06	0.66	0.02	0.06
100w_a1	5.88	0.00	0.00	0.63	0.49	0.09	3.33	0.13	0.17	0.05	0.11	0.57	0.06	0.02	0.13	0.00	0.11
100w_a2	5.67	0.02	0.00	0.50	0.52	0.20	2.79	0.04	0.12	0.11	0.11	0.65	0.00	0.07	0.43	0.11	0.00
100w_a3	5.96	0.07	0.00	0.53	0.41	0.23	3.06	0.24	0.22	0.21	0.10	0.47	0.02	0.02	0.16	0.12	0.08

100w_t	5.75	0.03	0.00	1.14	0.47	0.09	2.19	0.11	0.17	0.07	0.15	0.30	0.22	0.07	0.52	0.10	0.11
100w_a	5.77	0.03	0.00	0.54	0.49	0.17	2.99	0.13	0.17	0.12	0.11	0.58	0.03	0.03	0.24	0.08	0.06
Diff	0.02	0.00	0.00	-0.60	0.02	0.08	0.80	0.02	0.00	0.05	-0.04	0.28	-0.19	-0.04	-0.28	-0.02	-0.05

X7 1.14 → 0.54 (27人中4人が0になった)

X4b 0.52 → 0.24 (27人中6人が0になった)

# 減りにくいエラー

## X3(誤訳)、X13(句読法違反)

Total	Total	4a	6	7	1	2	3	8	10	11	12	13	9	16	4b	15	14
100w_t1	5.19	0.00	0.00	0.95	0.28	0.08	2.50	0.22	0.12	0.04	0.06	0.10	0.36	0.04	0.34	0.06	0.04
100w_t2	5.81	0.02	0.00	1.11	0.56	0.00	1.82	0.06	0.26	0.10	0.21	0.31	0.25	0.11	0.57	0.21	0.22
100w_t3	6.25	0.08	0.00	1.37	0.56	0.19	2.27	0.06	0.13	0.06	0.18	0.49	0.06	0.06	0.66	0.02	0.06
100w_a1	5.88	0.00	0.00	0.63	0.49	0.09	3.33	0.13	0.17	0.05	0.11	0.57	0.06	0.02	0.13	0.00	0.11
100w_a2	5.67	0.02	0.00	0.50	0.52	0.20	2.79	0.04	0.12	0.11	0.11	0.65	0.00	0.07	0.43	0.11	0.00
100w_a3	5.96	0.07	0.00	0.53	0.41	0.23	3.06	0.24	0.22	0.21	0.10	0.47	0.02	0.02	0.16	0.12	0.08

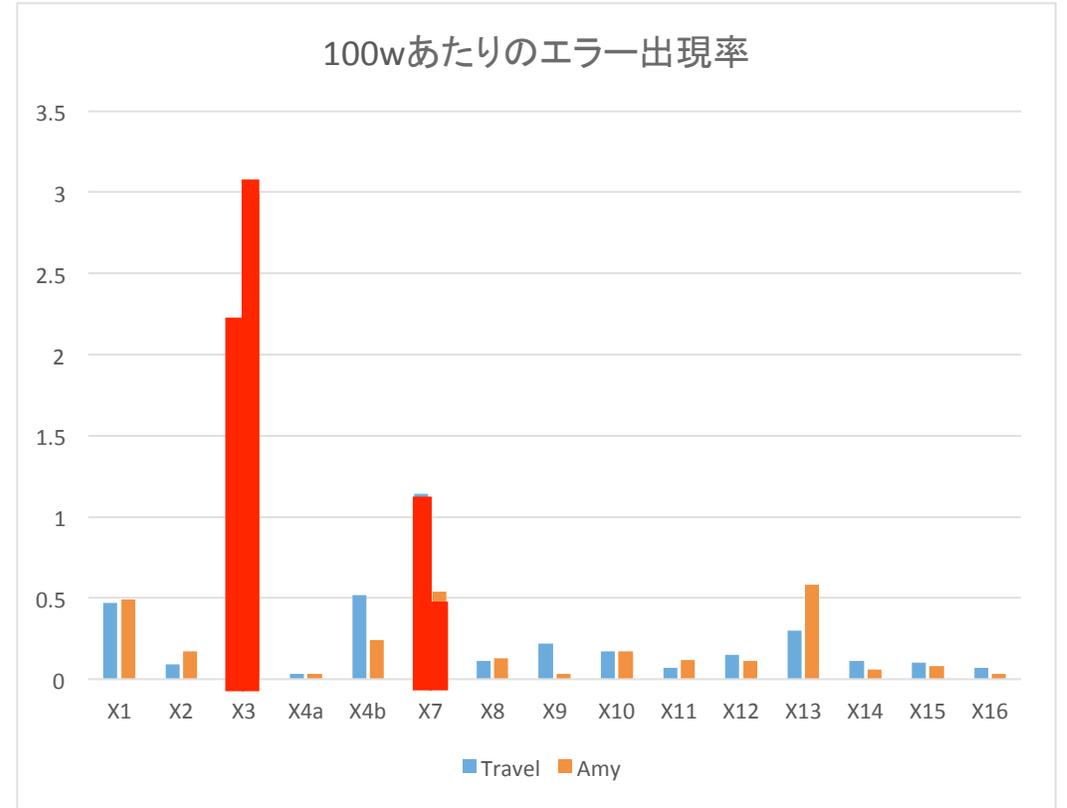
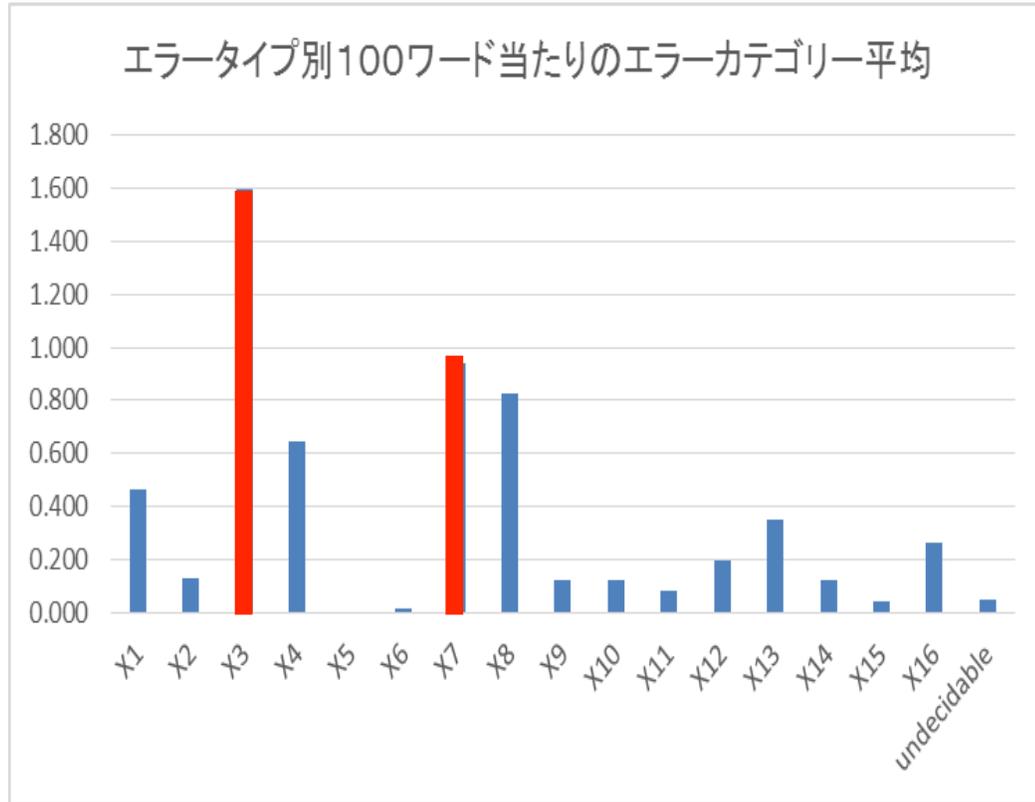
100w_t	5.75	0.03	0.00	1.14	0.47	0.09	2.19	0.11	0.17	0.07	0.15	0.30	0.22	0.07	0.52	0.10	0.11
100w_a	5.77	0.03	0.00	0.54	0.49	0.17	2.99	0.13	0.17	0.12	0.11	0.58	0.03	0.03	0.24	0.08	0.06
Diff	0.02	0.00	0.00	-0.60	0.02	0.08	0.80	0.02	0.00	0.05	-0.04	0.28	-0.19	-0.04	-0.28	-0.02	-0.05

X3 2.19 → 2.99

X13 0.30 → 0.58

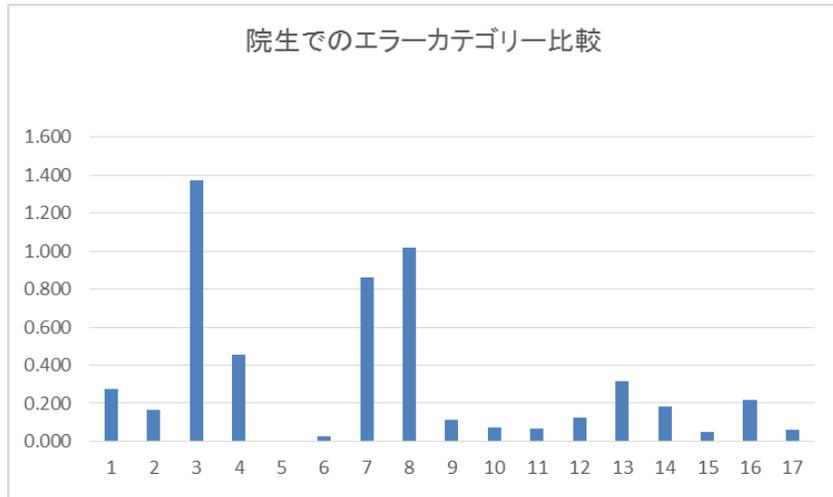
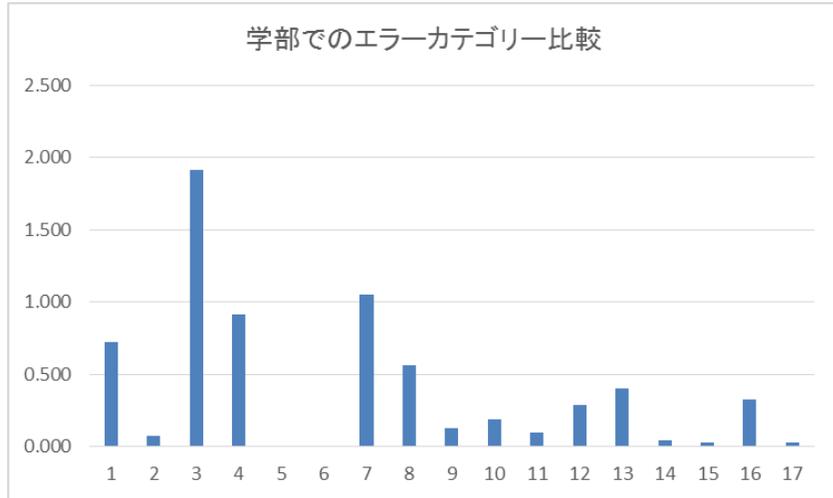
X3に関しては原文の難易度によるとも考えたが、課題ごとに見ても減っていない。

## b. Toyoshima (2014)との誤りの分布の比較

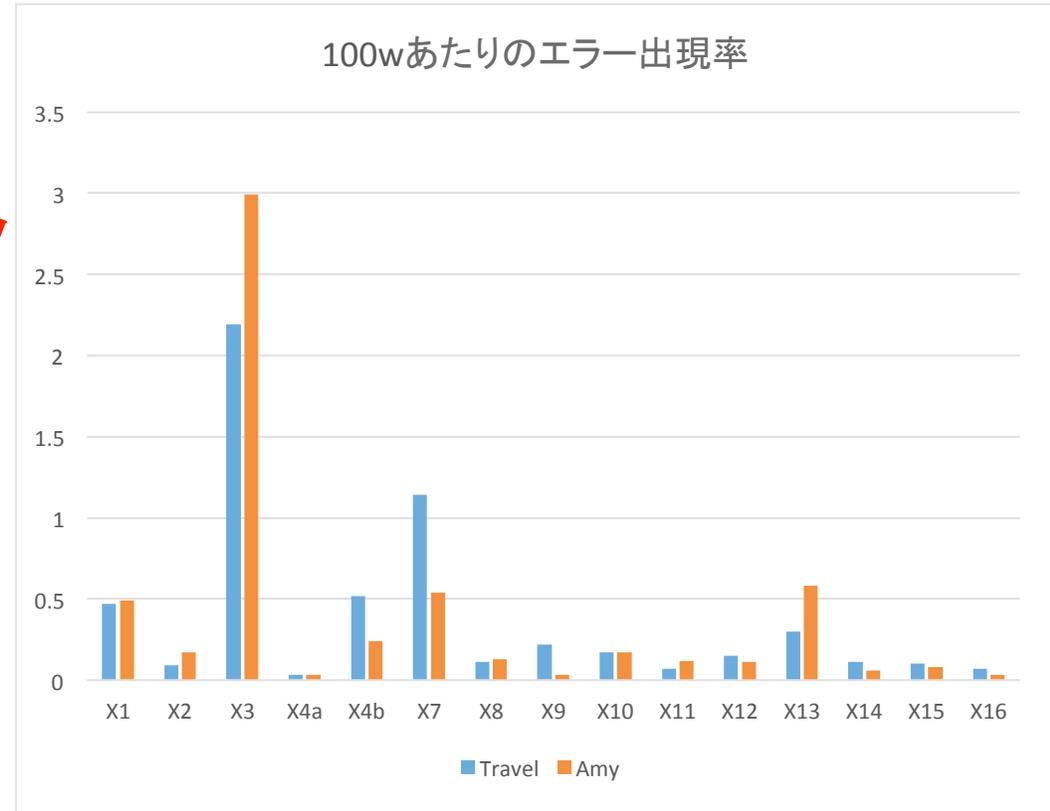


豊島 X3 X7 X8 X4(X4b) X1 X13 ...  
 今回 X3 X7 X1 X13 X4b ...

## b. Toyoshima (2014)との誤りの分布の比較(学部生と院生)



類似



## c. 文章産出困難感尺度(岸ら, 2012)を援用したアンケート調査

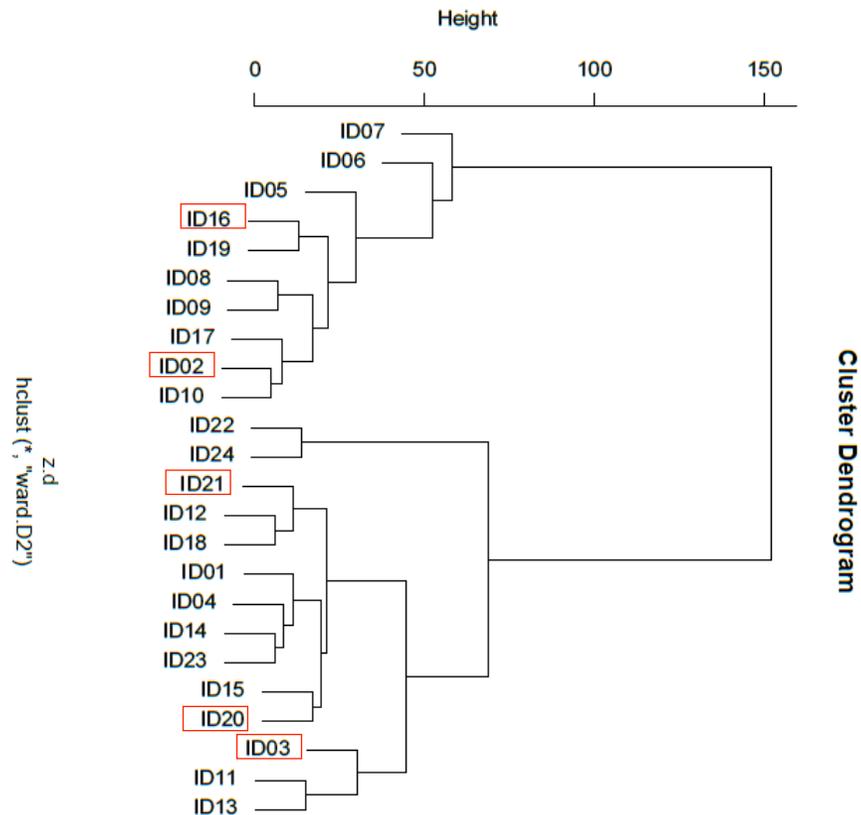
### • 意識化尺度(18問)

- 読み手に対して内容を正しく伝える
- 誤字・脱字を見つけて修正する
- 読み手の視点で文章を読み返す
- 読み手のことを考えて訳す                      など

### • 自己効力感尺度(6問)

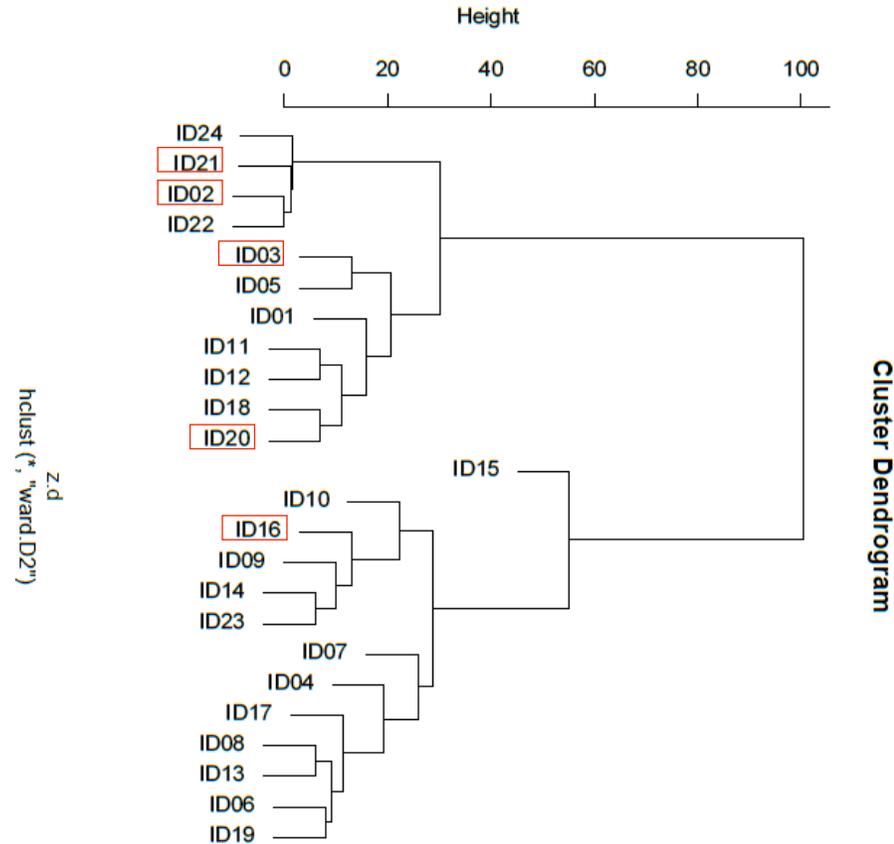
- 読み返したとき、自分の訳のどの部分を修正するとよいか分かる
- 読み返したとき、SDの内容が正しく訳せているかどうか分かる
- どの部分に、注意して読み返せばよいか分かる                      など

# 意識化に関する回答状況の類似性と X7とX4bの少なくとも一方が1.0以上減少した学習者



初回

意識が似てきた



最終回

- エラーカテゴリーの紹介
- 決定木の紹介
- 今回の研究内容
- 結果・考察
- **まとめ・今後**

## まとめ

- X7(一貫しない訳語)やX4b(直訳)は、回を重ねるごとに減りやすい。
- X3(誤訳)は原文レベルに関係なく、減りにくい。
- X13(句読法違反)は、毎回の指示があったにもかかわらず減りにくい。  
→細かい指定がややこしい? 意識の問題?
- Toyoshima (2014)とエラーの出現傾向が類似。
- Toyoshima (2014)の院生の出現エラーよりも、学部生の出現エラーと類似。  
→学部生のエラー傾向?
- エラーの出現率が大きく減った学習者は、  
翻訳に対する近い意識をもつようになった。

## 今後の研究計画

- エラーカテゴリーの4種類(内容・語彙表現・文法・テキスト)ごとにエラー推移を観察する。  
→エラーごとではなく、どんな種類のエラーが多いのか(減りやすいのか)などを調べる。
- 学生ごとのエラーの差を調べ、各エラーの増減と学生の傾向を見る。  
→人によってエラー出現傾向が違うことを調べる。

## 参考文献

- Bogdan Babych, Anthony Hartley, Kyo Kageura, Martin Thomas, Masao Utiyama (2012) “MNH-TT: a collaborative platform for translator training.” *Translating and the Computer* 34. 18 pages.
- Sara Castagnoli, Dragos Ciobanu, Kerstin Kunz, Natalie Kübler, Alexandra Volanschi (2006) “Designing a Learner Translator Corpus for Training.” In *Proc. of the 7<sup>th</sup> International Conference on Teaching and Language Corpora*.
- Anthony Hartley, 影浦峽, Martin Thomas, 内山将夫 (2014) 共同翻訳を考慮した「翻訳教育用みんなの翻訳」システム～みんなの翻訳第4報～. *言語処理学会第20回年次大会発表論文集*, pp.254-257.
- 岸学, 梶井芳明, 飯島里美 (2012) 文章産出困難感尺度の作成とその妥当性の検討. *東京学芸紀要*, Vol. 63, No. 1, pp. 159-169.
- Chiho Toyoshima (2014) “An analysis of error categories in learners of English-to-Japanese translation.” *Graduation thesis, Kobe College. Graduate Division of English <BB01258449>*.
- Chiho Toyoshima, Kikuko Tanabe, Anthony Hartley, Kyo Kageura (2015) “Error categories in English to Japanese translations.” *言語処理学会第21回年次大会発表論文集*, pp. 1076-1079.
- Chiho Toyoshima, Kikuko Tanabe, Atsushi Fujita, Anthony Hartley, Kyo Kageura (forthcoming)