

Error Types of Japanese Learners Corpus of Mandarin Chinese and NTNU/TUFS Online Dictionary of Misused Chinese

福田翔 林艶 申亜敏 游韋倫 佐野洋 望月圭子

東京外国語大学国際シンポジウム
Language Education and Computer Science for Second Language
Acquisition, e-Learning and Learners' Corpora

March 1, 2014, TUFS

Online Dictionary of Misused Chinese

Corpus features

- Making the corpus public on the internet
- Being able to search for sentences and words using different error classifications, the student's background, the topic of essay etc
- Being able to refer to the original essays and the revised essays

Corpus samples

18	...父母打电话, 回老家的时间, 我竟	想法 留心	也是孝敬父母的一个好方法, 第三...
学习者ID	該当部分前方	該当部分	該当部分後方

1 to 10 (21)

カテゴリから検索

复句の固定句式【一…就】
 复句
 名词
 副词
 动词
 主语
 对象介词
 原因介词
 形容词
 指示代词
 因果复句：因为……所以
 动态助词
 结构助词
 构造助词
 并列复句：既…也
 否定副词
 不及物动词
 时间名词
 固定句式：无论…也（都）…
 能愿动词
 趋向补语
 能愿动词
 趋向补语
 连词
 【表现】动词：和…商量
 表现
 因果复句：既然…就
 介词短语：在…里
 【表现】动词：从…开始

学習者情報から検索

18	...父母打电话, 回老家的时间, 我竟	想法 留心	也是孝敬父母的一个好方法, 第三...
学习者ID	該当部分前方	該当部分	該当部分後方

1 to 10 (21)

カテゴリから検索

学習者情報から検索

TOEFL (BT) : 0点 - 120点

学習期間 : 0年 - 20年

海外滞在経験 : あり なし

性別 : 男 女

年齢 : 10代 20代 それ以上

国籍 : 日本 中国 その他

母語 : 日本語 中国語 英語 その他

教育言語 : 日本語 中国語 英語 その他

中国語学習者中国語作文誤用コーパス検索サイト

サイトについて

本サイトは、中国語学習者コーパスのデータを検索できるWebアプリケーションです。コーパスデータは、東京外国語大学・中国語専攻・望月圭子研究室、佐野洋研究室、申重敏、福田翔岡 氏らの協力のもと収集されました。検索可能なデータ件数は、平成25年10月0日時点で〇〇件です。

誤用文字列 : ⇒ 修正文字列 :

タイプ検索 : 置換 削除 追加

1 to 10 (21)

学习者ID	該当部分前方	該当部分	該当部分後方
18	...東京, 东京的生活费用比我老家太	高昂	为减轻那个负担, 我还在打工挣钱...
18	...我老家贵, 所以父母的负担很大,	为了	减轻那个负担, 我还在打工挣钱, ...
18	...贵, 所以父母的负担很大, 为减轻	那个负担	负担, 我还在打工挣钱, 我挣的钱...
18	...自己打工挣一些钱, 我的负担	减小减少	了, "第二个是时间的问题, 上次...
18	...每天很忙, 没有时间, 所以我没有	时间	父母打电话, 或者去老家的时间, ...
18	...间, 所以我没有对父母打电话,	或者也	去老家的时间, 但是, 毕业大学, ...
18	...间, 所以我没有对父母打电话或者	去回	老家的时间, 但是, 毕业大学, 真...
18	...以, 我还在改变了自己想法, 无论	如何怎么	忙, 也都对父母打电话回老家的时...
18	...变了自己想法, 无论如何忙, 也挤	时间	父母打电话, 回老家的时间, 我觉...
18	...父母打电话, 回老家的时间, 我竟	想法 留心	也是孝敬父母的一个好方法, 第三...

Structure of the Talk

Corpus

- Explaining the students' essays
- Promissory letter and students' background
- How to correct essays
- How to set up and add error tags

Case study

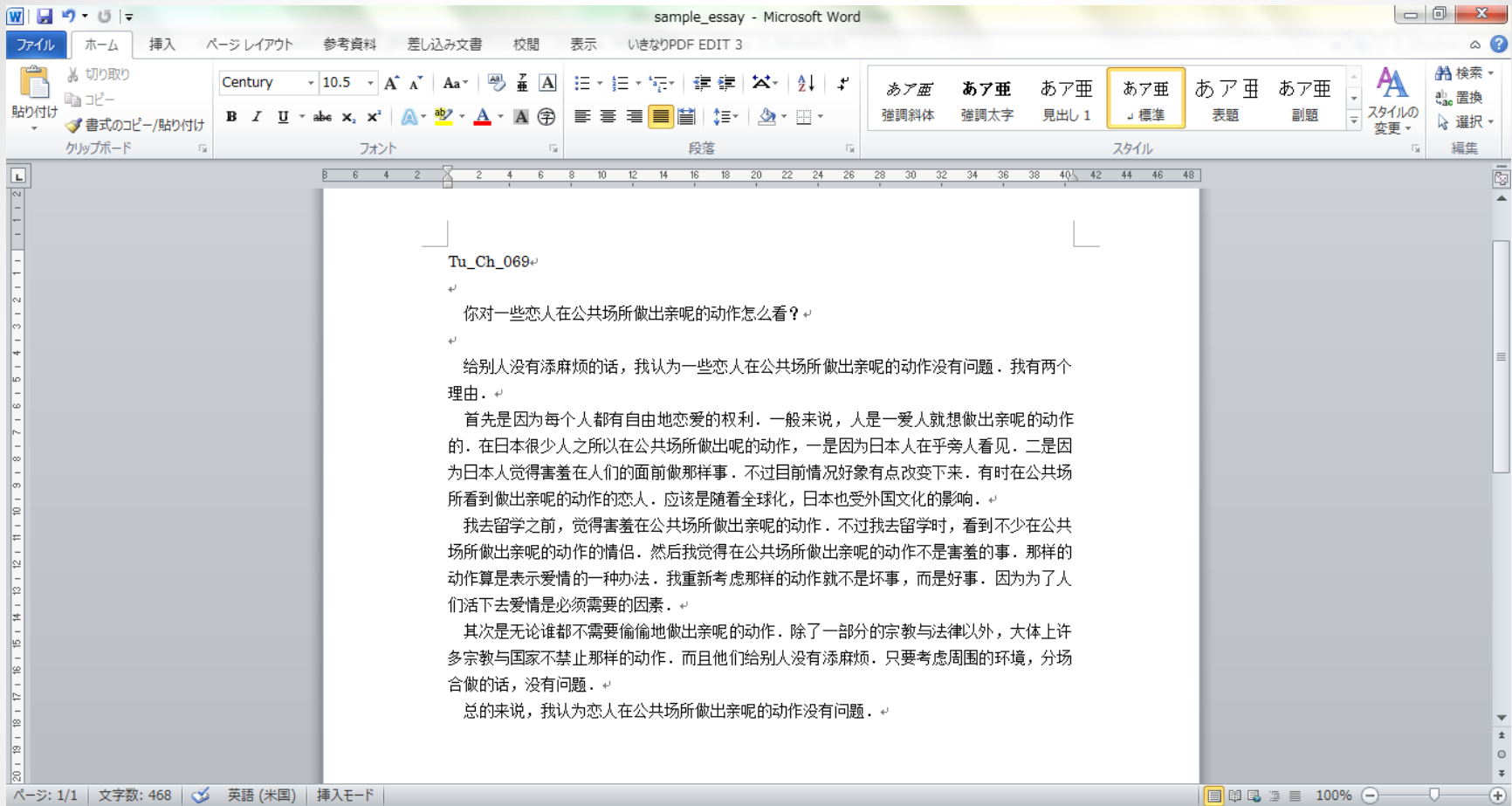
(Analysis of learners' misuse of Chinese complements)

- The six types of Chinese complements
- Frequency of correct and wrong use
- Error classification for each complement

Summary of Essay

- Essays are written by students of the Chinese department at TUFS (second to fourth year students).
- The total number of students who are taking part in this project is 81: 38 second year students, 26 third year students, 17 fourth students.
- Currently, we have 248 essays.
- Presently, we have a collection of 92,000 characters.

Essay Sample



- Student ID: Tu_Ch_069
- Format: Word (.docx)

Students' Background

Table1-1 Students' Background: Basic Information

ID	major	grade	name	sex	age	history of residence			nationality	mother tongue	language of education	history of studying chinese	institution of studying Chinese		experience of studying abroad		
						1	2	3					1	2	term	institution	location
Tu_Ch_001	Chinese	4	---	male	22	0-20Japan			Japan	Japanese	Japanese	4 years	TUFS		2011.9-2012.6	QingHua University	Pekin
Tu_Ch_002	Chinese	3	---	female	21	0-14China	14-21Japan		Japan	Chinese	Chinese/Japanese		TUFS				
.
.
.

Table1-2 Students' Background: Language

Language spoken at home			Language spoken with another person	Educational language of elementary school	Educational language of junior high school	Educational language of high school
with father	with mother	with borther				
Japanese	Japanese	Japanese	Japanese	Japanese	Japanese	Japanese
Japanese	Chinese		Japanese	Chinese	Japanese	Japanese
▪	▪	▪	▪	▪	▪	▪
▪	▪	▪	▪	▪	▪	▪
▪	▪	▪	▪	▪	▪	▪

Table1-3 Students' Background: Examination

EIKEN		TOEFL(iBT)						TOEIC				IELTS(academic)						IELTS(general)						Testing chinese proficiency		HSK (writing)		HSK(speaking)		
Grade	Date	R	W	S	L	Overall	date	R	L	Overall	Date	R	W	S	L	Overall	Date	R	W	S	L	Overall	Date	Grade	Date	Grade	Date	Grade	Date	
2	2007.2									745	2012.2																6	2012.1		
																								3	2023.6					
.	
.		
.		

Correction

Sample of Corrected Essay

Tu_Ch_001

第2课表达练习 3-3

你认为应该怎样孝敬父母？

我认为孝敬父母时，我要表示三个方面的感谢。第一个方面是，关于金钱方面的感谢。我老家在群馬，不过现在我住在東京。東京的生活費用比我老家太貴，所以父母的負擔很大。為了減輕那個他們的負擔，我現在打工掙錢。我掙的錢還沒不那麼多，不過我媽媽常常說，“因為你現在自己打工掙了一些錢，所以我們的負擔減少少。”第二個是關於時間的方面的感謝。上大學以後，我的生活很充實，每天很忙，沒有時間。所以我既沒有對給父母打電話，或者也沒有回去老家的時間。

コメント [潘55]: 【删除】【主语】

コメント [S56]: 动词

コメント [S57]: 介词 (对象)

コメント [m58]: 【替换】高很多【程度补语】

コメント [S59]: 介词 (原因)

コメント [潘60]: 【替换】他们【指示代词】

コメント [潘61]: 【添加】的【结构助词】

コメント [S62]: 代词

コメント [潘63]: 【删除】【否定副词】

Word format, comment function

Tag

- **Type of Errors**

「删除」(removal), 「添加」(addition), 「替换」(replacement), 「移动」(transfer)

- **Category of Grammar**

张斌 (2007) 《新编现代汉语》 复旦大学出版社.

齐沪扬 (2012) 《现代汉语》 商务印书社.

齐沪扬 (2010) 《对外汉语教学语法》 复旦大学出版社.

- Sample of Tag List

category	subcategory
名词	专有名词 可量名词 不可量名词 时间名词 处所名词 方位词
可量名词	个体名词 集体名词 物质名词 抽象名词
数词	系数词 位数词 系位数词 数量数词 概数词
动词1	动作动词 存现动词 关系动词 能愿动词 趋向动词 使令动词
动词2	及物动词 不及物动词
动词3	静态动词 持续动词 瞬间动词 终结动词
动词4	重叠动词
形容词	性质形容词 状态形容词
区别词	区别词
副词	程度副词 范围副词 时间副词 情态副词 否定副词 语气副词 关联副词
代词	人称代词 指示代词 疑问代词
连词	
介词	时间介词 处所介词 方向介词 范围介词 工具介词 依据介词 对待介词 原因介词 目的介词 与事介词 处置介词 被动介词 介引介词 伴随介词 比较介词
助词	结构助词 时态助词 时制助词 比况助词 表数助词 列举助词 其他助词
倒装句	主谓倒装 状语后置 宾语前置
句式	主谓谓语句 “把”字句 “被”字句 连动句 兼语句 存现句 比较句 “连”字句
联合复句	并列复句 承接复句 递进复句 选择复句 注解复句
偏正复句	因果复句 条件复句 转折复句 让步复句 目的复句

• Sample of Adding Tag System

中国語読用コーパスタク付ソフト (東京外大望月生子科併専用)

フォルダ tufs中文系 ファイル 3_Tu_Ch_039/Task/中国語2年生/学習歴1年1ヶ月/ 読込 保存 XML化

元に戻す やり直し タグ 検索 全検索 全タグ検査

水曜2限

我們找工作的時候，公司的人既<刪除 / 否定副詞 / 不→〇>看<替換 / 表現 / 名詞 / 表面的性格→外表>，又看<替換 / 表現 / 名詞 / 內面的性格→內在>。我們以前過<添加 / 結構助詞 / 情態補語 / 〇→得>怎麼樣，學到什麼，<替換 / 正反問句 / 有沒有→具不具備>在公司做<添加 / 結構助詞 / 情態補語 / 〇→得>好的能力，這些事情很重要。公司<刪除 / 表現 / 名詞 / 人→〇>評價我們以前的人生，還有現在的我們。

所以我覺得六個因素都很重要。(①有邏輯性②挑戰極限③明確目標④開闊眼界⑤<替換 / 錯別字 / 又→有>協調性⑥有獨立思考<添加 / 結構助詞 / 偏正短語 / 〇→的>能力) 如果我<替換 / 動詞 / 有→具備>這六個因素，對公司<刪除 / 表現 / 名詞 / 人→〇>來說，我<添加 / 語氣副詞 / 〇→就>是<添加 / 量詞短語 / 〇→一個>很有用的人。

這六個因素之中，<替換 / 時間副詞 / 連詞 / 如果→首先>我們<替換 / 能願動詞 / 可以→應該>開闊眼界，<添加 / 連詞 / 因果複句 / 〇→因為>開闊眼界<添加 / 能願動詞 / 〇→會>影響<替換 / 偏正短語 / 取得別的因素→別的因素的取得>。開闊眼界後，我們可以看得<添加 / 程度副詞 / 〇→更>清楚別人<添加 / 時間副詞 / 〇→在>做什麼，有人比我們自己更努力，這些事<添加 / 範圍副詞 / 〇→都><添加 / 能願動詞 / 〇→會>刺激我們的積極性，<替換 / 使令動詞 / 表現 / 叫→讓>我們挑戰極限。

而且，開闊眼界<添加 / 能願動詞 / 〇→可以><添加 / 使令動詞 / 兼語句 / 〇→讓我們>看到各種各樣的世界<替換 / 動詞 / 導向→引導><替換 / 人稱代詞 / 自己→自我>的形成，所以開闊眼界對<替換 / 動詞 / 有→培養>獨立思考能力<添加 / 程度副詞 / 〇→很>有用。

開闊眼界對<添加 / 動詞 / 有→培養>邏輯性也很有用，因為<添加 / 能願動詞 / 〇→要><替換 / 動詞 / 表現 / 知道→了解><替換 / 程度副詞 / 很→更>多<添加 / 結構助詞 / 〇→的>世界，邏輯性<添加 / 程度副詞 / 〇→很>有效。<添加 / 連詞 / 轉折複句 / 〇→但是>即使有邏輯性，沒開闊視野的話，因為<替換 / 表現 / 有很少的烤爐因素→考慮的因素太少>，邏輯性<添加 / 關聯副詞 / 〇→也><添加 / 關聯副詞 / 〇→就>沒有效<添加 / 語氣詞 / 〇→了>。

所以，這六個<替換 / 錯別字 / 音素→因素>之中開闊眼界特別重要。

折り返し 文字サイズ

- Example

(1) 所以開闊眼界對〈替換 / 动词 / 有→培养〉

獨立思考能力〈添加 / 程度副词 / ○→很〉有用。

Case Study: Analysis of Complements

Complement is bound to and follows the verb and expresses the result of the action of the verb.

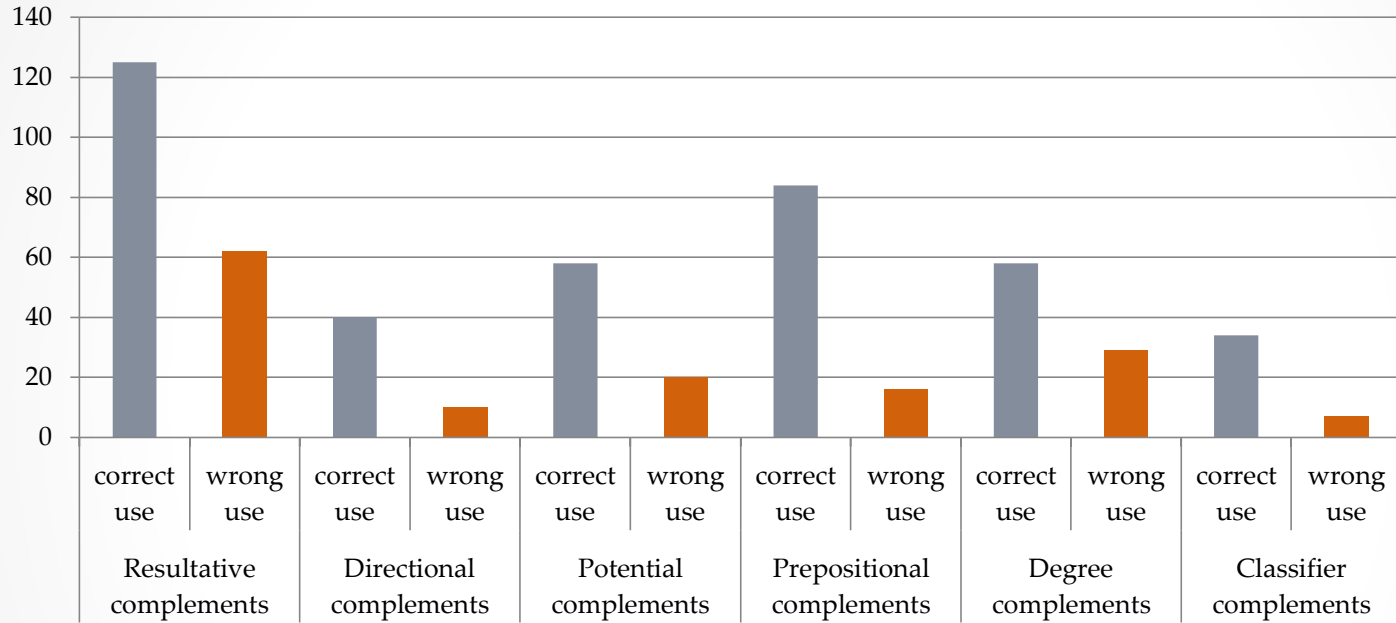
Five types of Chinese Complements

- Resultative Complement (结果补语)
- Directional Complement (趋向补语)
- Potential Complement (可能补语)
- Degree Complement (程度/情态补语)
- Classifier Complement (动量/数量补语)

Data

- The amount of essays: 100
- The number of characters: about 46,000

Rate of Complement Use



The Percentage of Correct and Wrong Use

type of complements	Resultative		Directional		Potential		Prepositional		Degree		Classifier	
type of use	correct	wrong	correct	wrong	correct	wrong	correct	wrong	correct	wrong	correct	wrong
%	67%	33%	80%	20%	74%	26%	84%	16%	67%	33%	83%	17%

Resultative Complement

(2a) *农村人不容易买 ϕ 好东西。

(2b) 农村人不容易买到好东西。

“It is difficult to purchase in the village.”

(3a) *你听 ϕ 意大利菜，想起来什么？

(3b) 你听到意大利菜，想起来什么？

“When you hear Italian food, what do you think of?”

- The most common error is to drop the complement “到”.
- The error type amounts to 68.3 % of the total (43 cases / 63 cases)

(4) 从小开始学习语言的话，比较不会怕说错，自由自在地讲。

“When you start to study a language from childhood, you don't worry about what you say is wrong and speak freely.”

Directional Complement

(5a) *原来这个节日是从中国 ϕ 来的风俗习惯，现在成为日本的一个重要节日。

(5b) 原来这个节日是从中国传来的风俗习惯，现在成为日本的一个重要节日。

“This holiday originally came from customs and habits, but now it has become an important holiday in Japan.”

(6a) *我们要为了自己 ϕ 出去社会好好地干。

(6b) 我们要为了自己走出社会好好地干。

“We have to go into the world and work hard for ourselves.”

- The most common error is to drop the preceding verb.

Potential Complement

(7a) *但就算是多么高价的礼物胜不了子女的体贴入微。

(7b) 但就算是多么高价的礼物也比不上子女的体贴入微。

“But such an expensive present can not be compared with the warmth felt by children.”

- The most common error is the use of vocabulary.

Degree Complement

(8a) *女性不用被家务拘束，我们要为了自己走出社会干得好。

(8b) 女性不用被家务拘束，我们要为了自己走出社会好好地干。

“Women need not be restricted to housework. We can go into the world and work hard for ourselves.”

(9a) *农村的话，孩子们可以去山里或草原玩儿得很痛快。

(9b) 农村的话，孩子们可以去山里或草原痛快地玩儿。

“In a village, children can play vigorously in the mountains and fields.”

- The most common error is not to use an appropriate phrase form.

Classifier Complement

(10a) *到目前为止，我三次去过中国 。

(10b) 到目前为止，我去过三次中国。

“Presently, I have been to China three times.”

- The most remarkable error is the incorrect word order.

The Most Remarkable Error for Each Complements

Kinds of complement	The causes of errors
Resultative Complement	to drop the phase complement
Directional Complement	to drop the preceding verb
Potential Complement	the use of vocabulary
Degree Complement	not to use an appropriate phrase form
Classifier Complement	the incorrect word order