

# Post-editing of Structurally Ambiguous Translation: The biasing effect from source text

Yuxiang Wei

Centre for Translation and Textual Studies  
Dublin City University, Ireland

yuxiang.wei3@mail.dcu.ie



Translating and the Computer (TC42)

20 November 2020

- Lexical and structural ambiguity
  - Track the influence of linguistic variables on the effort of post-editing
  - Understand the cognitive processes underlying the behaviour
- 
- This presentation:
    - Lexical disambiguation (Wei, forthcoming)
      - Word translation entropy (Carl et al. 2016)
    - Structural disambiguation
      - Product
      - Process

# Translation behaviour: affected by ambiguity

- Number of translation alternatives
- Eye-key span (Dragsted 2010; Dragsted & Hansen 2008)
  - Carl & Schaeffer (2017): Larger sample and more rigorous testing
  - Danish, Spanish, Chinese, Hindi, German
  - 108 participants; 3,242 unique source words
- Decontextualised single words
  - (Tokowicz & Kroll 2007)
- Psycholinguistics: non-selective activation of both languages (Brysbaert 1998; DeBrujin et al. 2001; Dijkstra et al. 2000; Marian et al. 2003; Schwartz & Kroll 2006)

# Word Translation Entropy

- Degree of uncertainty
- Better measure for the variation of translation alternatives (Bangalore et al. 2016)
- Co-activated translation options
- Significant impact on different measures of effort (e.g. Schaeffer et al. 2016)
  - First fixation duration
  - Total reading time
  - Word production duration
  - Probability of a fixation

# Translation ambiguity (lexical)

- Conceptual basis on which entropy can describe the cognitive activities, and represent the cognitive load, in translation
- Entropy values of the words fixated in activity units
- Cognitive processing of a high-entropy item is facilitated by the contextual information provided by items of low entropy levels

Wei, Y. (forthcoming). Entropy and eye movement: A micro analysis of information processing in activity units during the translation process. In Michael Carl (ed): Explorations in Empirical Translation Process Research

HTra: 2.0357

STid: 4 5 6 7 8 9 10 11 12 13 14 15 16  
ST: ...increase in cost of living British families have to cough up an extra  
17 18 19 20  
£ 31,300 a year...

Path:  
12 13 13 10 5 5 4  
to cough cough families in in increase

Htra / HCross: cough – 3.3755 / 1.2087  
to – 0.7654 / 0.2580  
families – 0.2580 / 0.2580  
in – 2.0185 / 0.5132  
increase – 2.4556 / 0.4262

# Structural ambiguity

- Data
  - CRITT TPR-DB
- Language
  - English to Chinese
- Task
  - Post-editing
- Participants
  - Native Chinese speakers, English as a second language
  - with 1-year translation experience
  - 36 -> 18
- Structural ambiguity in MT output (garden path)

# Garden path (temporary ambiguity)

- When the dog scratched the vet and his new assistant removed the muzzle.
- *(Main verb / Reduced relative clause)*
  - The horse raced past the barn.
  - The horse raced past the barn fell.
- *(Noun phrase / Sentential complement)*
  - The wife will claim the inheritance.
  - The wife will claim the inheritance belongs to her.
  - The wife will claim the husband has left (的/de) inheritance.
    - A- main verb                      B-verb                      C- head noun
- Single/multiple interpretation, thematic role assignment, argument integration



# MT Sentence

- 我们提出我们主要围绕关键的全球优先事项和当前世界面临的挑战，  
如粮食安全和能源领域的科研实力。

- A-提出 (main verb)

1a: Sentential complement and preposition

- B-围绕 (v. /prep. )

1b: Sentential complement and verb

- C-科研实力 (head noun)

2a: Noun phrase and preposition

2b: Noun phrase and verb

- 我们 [提出] 我们主要 [围绕] 关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的 [科研实力]。

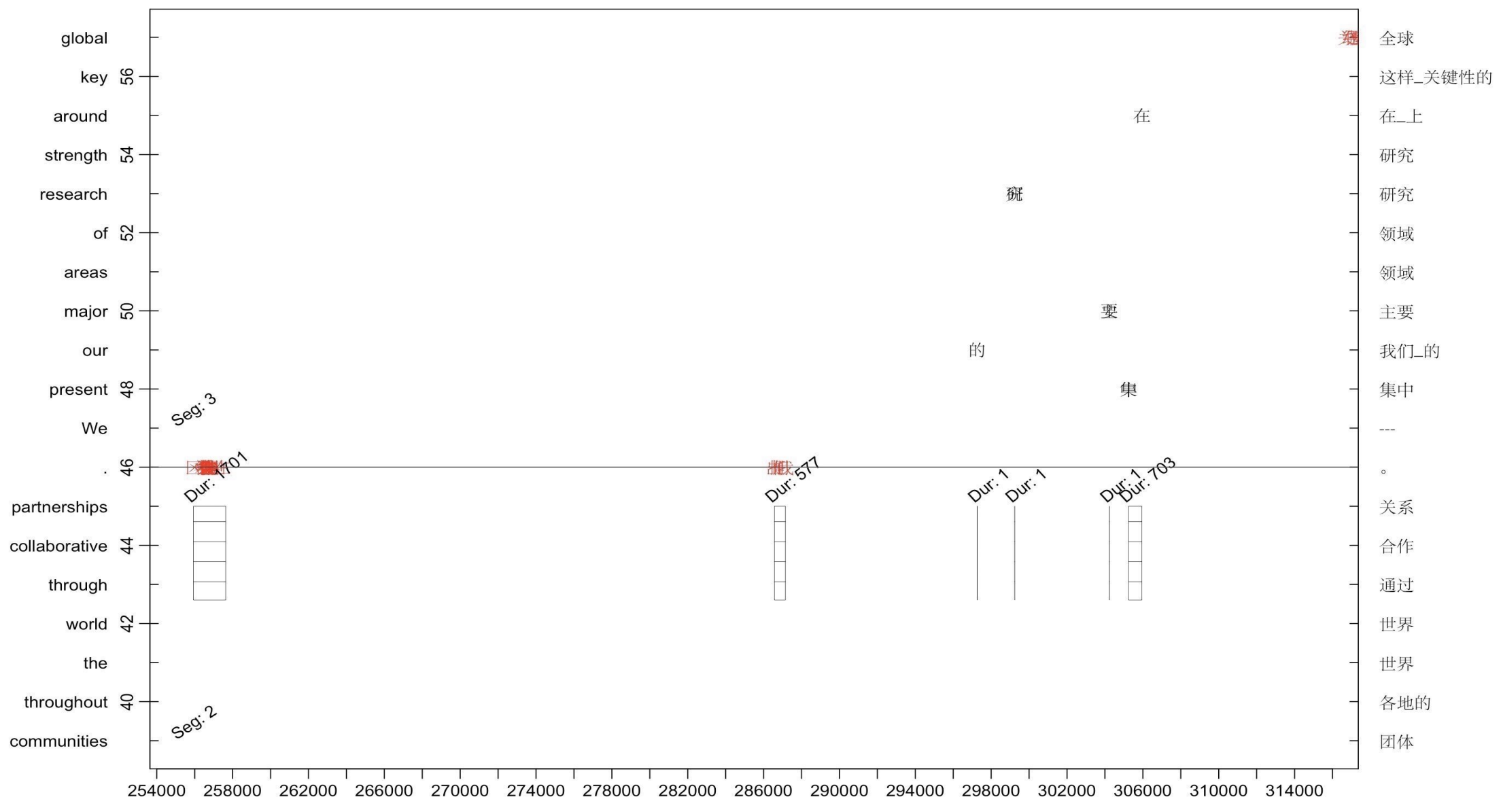
Wǒmen [ tíchū ] wǒmen zhǔyào [ wéirào ] guānjiàn de quánqiú yōuxiān shìxiàng hé dāngqián shìjiè miànlín de tiǎozhàn, rú liángshí ānquán hé néngyuán lǐngyù de [ kēyán shílì ].

We [ propose ] that we mainly [ focus on ] key global priorities and challenges currently facing the world, such as [ research capabilities ] in the fields of food security and energy.

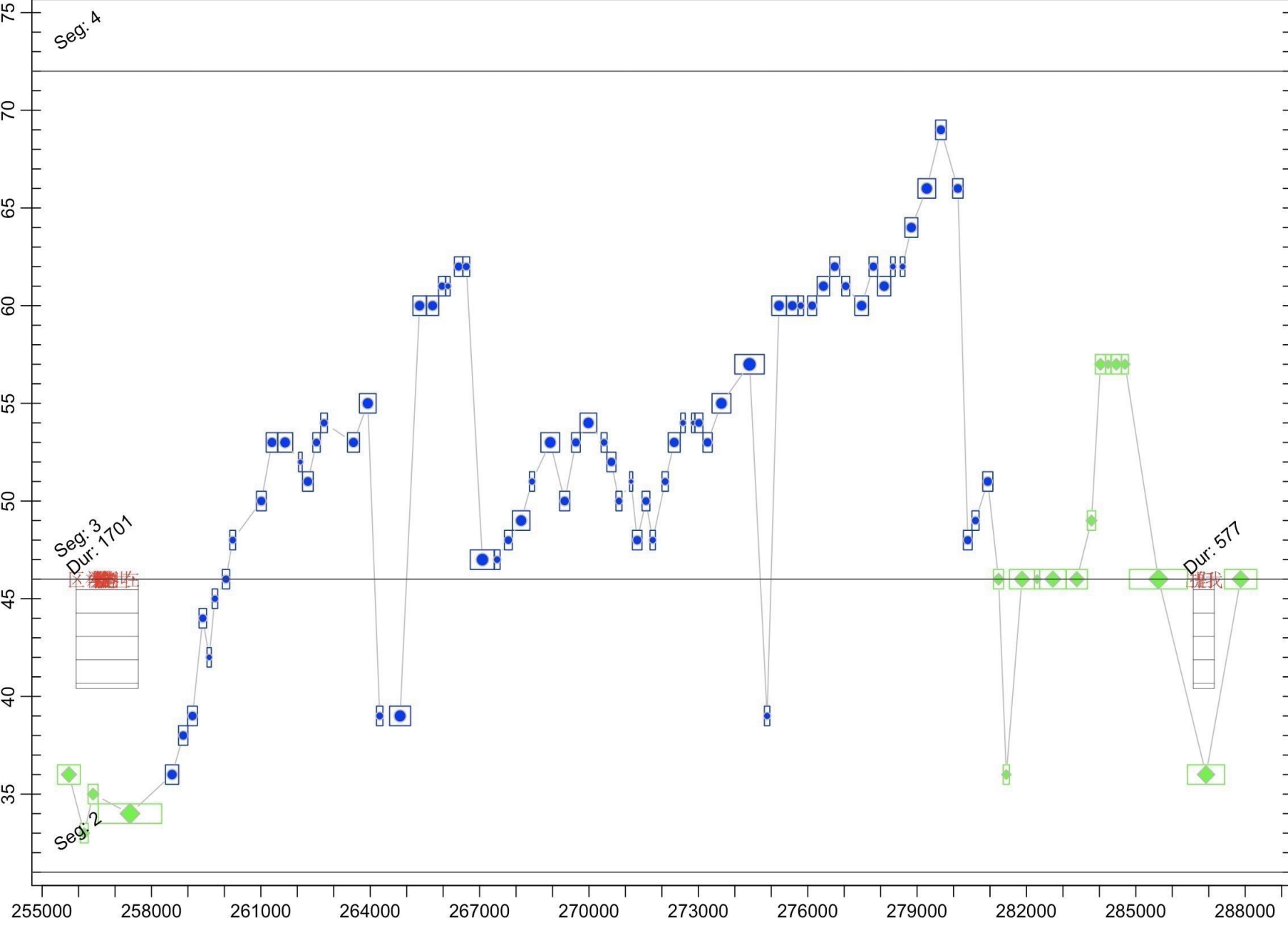
We [ present ] our major areas of [ research strength ] [ around ] key global priorities and challenges currently confronting the world, such as food security and energy.

- A-提出 (main verb)      tíchū      to propose, put forward, raise, present
- B-围绕 (v. /prep. )      wéirào      around / to surround
- C-科研实力 (head noun)      kēyán shílì      research strengths, research capabilities

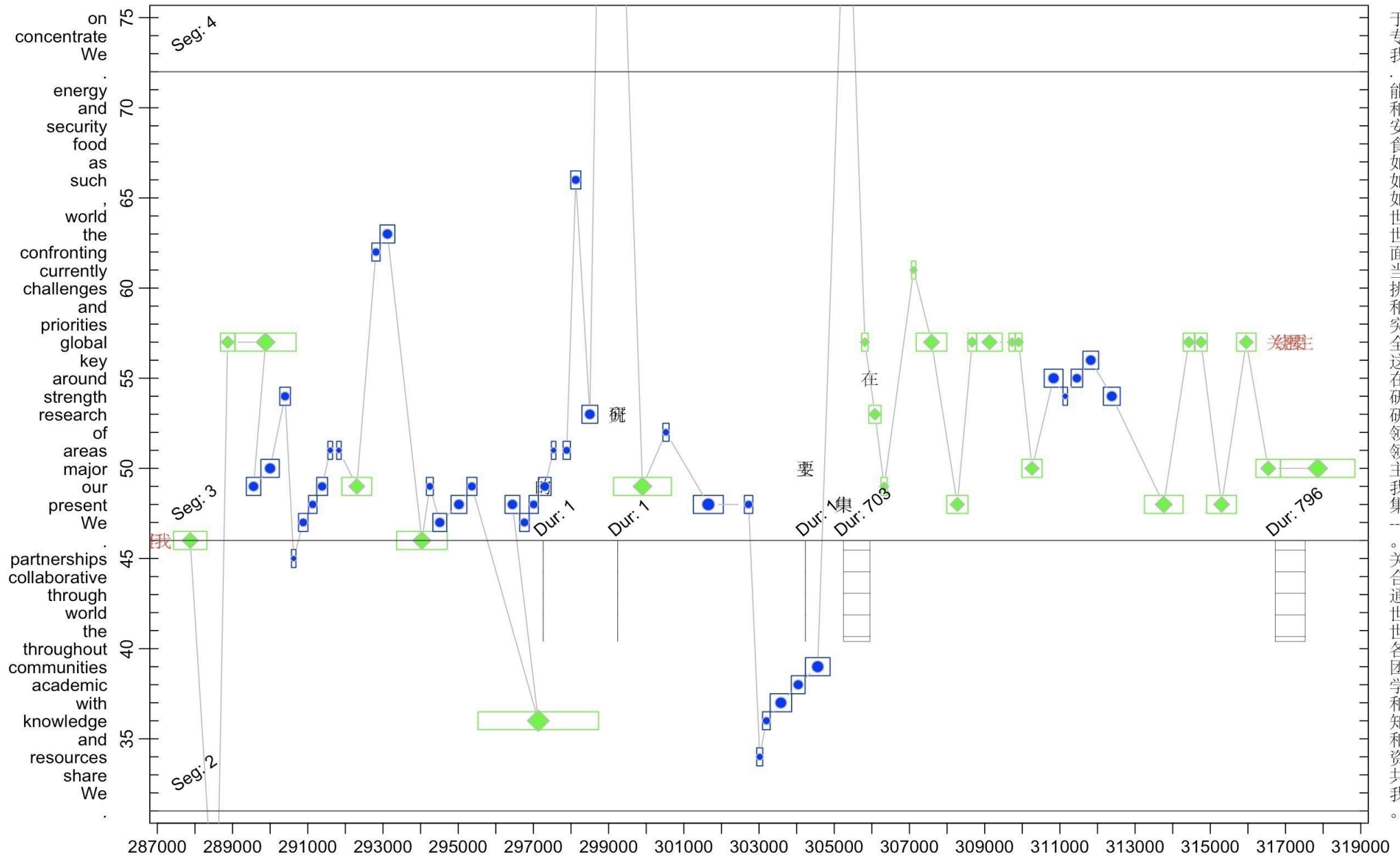
# Post-editing process



on  
concentrate  
We  
energy  
and  
security  
food  
as  
such  
world  
the  
confronting  
currently  
challenges  
and  
priorities  
global  
key  
around  
strength  
research  
of  
areas  
major  
our  
present  
We  
partnerships  
collaborative  
through  
world  
the  
throughout  
communities  
academic  
with  
knowledge  
and  
resources  
share  
We



于  
专注  
我们  
能源问题  
和安全  
食品  
如  
如  
世界  
世界  
面临的  
当前  
挑战  
和  
突出  
事项  
全球  
这样  
关键性的  
在上  
研究  
研究  
领域  
领域  
主要  
我们  
的  
集中  
---  
。关系  
合作  
通过  
世界  
各地的  
团体  
学术  
和  
知识  
和  
资源  
共享  
我们  
。



于  
专注  
我们  
·  
能源\_问题  
和  
安全  
食品  
如  
如  
世界  
世界  
面临的  
当前  
挑战  
和  
突出\_事项  
全球\_关键性的  
这样\_上  
在\_研究  
研究领域  
主要\_的  
集中  
---  
。关系  
合作  
通过  
世界  
各地的  
团体  
学术  
和  
知识  
资源  
共享  
我们  
。

在华威，我们一如既往地明确承诺,所有的学科都成为世界一流的研究和创新中心。我们和世界各地的学术团体通过合作关系共享资源和知识。我们提出我们主要围绕关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的科研实力。我们专注于华威社会研究，特别是在知识转移方面的积极影响。我们的目的是要进行令人兴奋的，突破性的，优秀的，在许多情况下，与政策相关的研究。

Position of A, B, and C

在华威，我们一如既往地明确承诺,所有的学科都成为世界一流的研究和创新中心。我们和世界各地的学术团体通过合作关系共享资源和知识。我们提出我们主要围绕关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的科研实力。我们专注于华威社会研究，特别是在知识转移方面的积极影响。我们的目的是要进行令人兴奋的，突破性的，优秀的，在许多情况下，与政策相关的研究。

Fixations 1-4

在华威，我们一如既往地明确承诺,所有的学科都成为世界一流的研究和创新中心。我们和世界各地的学术团体通过合作关系共享资源和知识。我们提出我们主要围绕关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的科研实力。我们专注于华威社会研究，特别是在知识转移方面的积极影响。我们的目的是要进行令人兴奋的，突破性的，优秀的，在许多情况下，与政策相关的研究。

Fixations 5-10

在华威，我们一如既往地明确承诺,所有的学科都成为世界一流的研究和创新中心。我们和世界各地的学术团体通过合作关系共享资源和知识。我们提出我们主要围绕关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的科研实力。我们专注于华威社会研究，特别是在知识转移方面的积极影响。我们的目的是要进行令人兴奋的，突破性的，优秀的，在许多情况下，与政策相关的研究。

Fixations 11-12



- Incremental processing: garden path, disconfirmed by C
- Head-driven approach: increased effort regarding B
- Krings (2001)
- Bilingual co-activation
  - Syntactic processing of the TT can co-occur with ST comprehension (Jensen et al., 2009; Balling et al., 2014)
  - Disambiguation and re-analysis on the MT sentence seems to be bypassed
- 我们提出我们主要围绕关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的科研实力。

# Post-edited output

- 我们提出我们主要围绕关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的科研实力。

1. B as a preposition attached to A (verb), C as object of A
2. B as a verb, with C as subject (converting the head-final phrase to a clause)

• Category	0	1	2a	2b	2c	2d
• Frequency	2	5	8	1	1	1
• Proportion	0.111	0.278	0.444	0.056	0.056	0.056

- Category 2
- 11 out of 16



- Next step: Statistical analysis
  - reading time & fixation counts
- Another two sentences
  - Structural ambiguity without garden path
  - Unambiguous structurally
  - (similar structure with the sentence above – head-final nominal structure)
- Ambiguity in MT output: likely to be overlooked by the post-editor?

# Thank you!

Yuxiang Wei

*yuxiang.wei3@mail.dcu.ie*



Centre for Translation and Textual Studies  
Dublin City University, Ireland



# Sentence

- 我们**提出**我们主要**围绕**关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的**科研实力**。

## 1a. Sentential complement and preposition

我们 提出

*women tichu*

We present/propose

[N V

我们 主要 围绕 ... 如...

*women zhuyao weirao ... ru...*

we mainly around ... such as...

N ADV PREP NP PP] [VP]?

# Sentence

- 我们**提出**我们主要**围绕**关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的**科研实力**。

## 1b. Sentential complement and verb

我们	提出	我们	主要	围绕	...	如...
women	ichu	women	zhuyao	weirao	...	ru...
We	present/propose	we	mainly	surround	...	such as...
[N	V	N	ADV	V	NP	PP]

# Sentence

- 我们**提出**我们主要**围绕**关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的**科研实力**。

## 2a. Noun phrase and preposition

我们	提出	我们	主要	围绕	...	的	科研实力
<i>women</i>	<i>tichu</i>	<i>women</i>	<i>zhuyao</i>	<i>weirao</i>	<i>...</i>	<i>de</i>	<i>keyanshili</i>
We	present/propose	our	mainly	around	...	<i>de</i>	research strength
[N	V	DET	ADV	PREP	NP	<i>de</i>	HN]

# Sentence

- 我们**提出**我们主要**围绕**关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的**科研实力**。

## 2b. Noun phrase and verb

我们	提出	我们	主要	围绕	...	的	科研实力
<i>women</i>	<i>tichu</i>	<i>women</i>	<i>zhuyao</i>	<i>weirao</i>	<i>...</i>	<i>de</i>	<i>keyanshili</i>
We	present/propose	we	mainly	surround	<i>...</i>	<i>de</i>	research strength
[N	V	N	ADV	V	NP	<i>de</i>	HN].

- 我们**提出**我们主要**围绕**关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的**科研实力**。

- 1a: Sentential complement and preposition
- 1b: Sentential complement and verb
- 2a: Noun phrase and preposition
- 2b: Noun phrase and verb



# Post-edited output

- 我们**提出**我们主要**围绕**关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的**科研实力**。

1. Word B as a preposition which is attached to word A:

a. 我们	围绕...	展现	... 科研实力
women	weirao...	zhanxian	... keyan shili
We	around...	present	... research strength
[N	PP	V	NP]

b. 围绕...	我们	展现	实力 (n)
weirao...	women	zhanxian	keyan shili
Around...	we	present	research strength
[PP	N	V	NP]

# Post-edited output

- 我们**提出**我们主要**围绕**关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的**科研实力**。

2. *Word B as a verb, with word C as subject (omitting word A, and converting the head-final phrase to a clause):*

我们的	科研实力	围绕...
<i>women de</i>	<i>keyan shili</i>	<i>weirao...</i>
Our	research strength	surround...
[DET	NP	VP]

3. *Merged with the next sentence, but adopting a similar structure to 2:*

我们	围绕	...	并...
<i>women</i>	<i>weirao</i>	...	<i>bing...</i>
<i>We</i>	<i>surround</i>	...	<i>and...</i>
[N	V	NP	Conjunction]

# Post-edited output

- 我们**提出**我们主要**围绕**关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的**科研实力**。

4. Adopting the original TT syntax for word A (using the sentential complement interpretation), while converting the head-final structure to a clause (i.e. as in 2):

我们	提出	优势领域	是围绕...
women	tichu	youshi lingyu	shi weirao...
We	present/propose	advantageous area	is around...
[N	V	NP	VP]

5. Converting the embedded structure into a paratactic one (where the first chunk contains word A as the verb and word C as the object, and the second chunk is the same as 2):

我们	显示	实力，	这些领域	关乎
women	xianshi	keyan shili,	zhexie lingyu	guanhu
We	present	research strength,	these areas	are about
[N	V	NP,	NP	VP]

- Lexical and structural ambiguity
- Track the influence of linguistic variables on the effort of post-editing
- Understand the cognitive processes underlying the behaviour
- Data:
  - CRITT TPR-DB
  - Eye-tracking and key-logging experiments
- This presentation:
  - Lexical disambiguation (Wei, forthcoming)
    - Word translation entropy (Carl et al. 2016)
  - Structural disambiguation
    - Product
    - Process

# Post-edited output

- 我们提出我们主要围绕关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的科研实力。
1. B as a preposition attached to A (verb)
  2. B as a verb, with C as subject (omitting A, and converting the head-final phrase to a clause)
  3. Merged with the next sentence, but adopting a similar structure to 2
  4. Adopting the original TT syntax for A (using the sentential complement interpretation), while converting the head-final structure to a clause (i.e. as in 2)
  5. Converting the embedded structure into a paratactic one (where the first segment contains A as the verb and C as the object, and the second segment is the same as 2)

# Post-edited output

- 我们提出我们主要围绕关键的全球优先事项和当前世界面临的挑战，如粮食安全和能源领域的科研实力。

• Category	0	1	2	3	4	5
• Frequency	2	5	8	1	1	1
• Proportion	0.111	0.278	0.444	0.056	0.056	0.056

- Categories 2, 3, 4, 5
- 11 out of 16