

Text Complexity & Translation Difficulty

Sanjun Sun

MC2 Lab Food for Thought series @ U of Bologna
2021-04-15



Outline

- Significance
- History
- Concepts
- Research framework
- Measurement methods
- Research objects
- Some findings

Significance: Theoretical

- In **Translation Studies**, among the core questions are:
 - ✓ What **difficulties** do translators (and interpreters) encounter in their work?
 - ✓ What strategies and methods do translators use to overcome these **difficulties**?
 - ✓ To what extent do translators solve these **difficulties**?



Significance: Theoretical

- In **Cognitive Translations Studies**:
- Muñoz's (2014) lists **seven topics or research areas**: competence and expertise, **mental load and linguistic complexity**, advances in research methods.....
- Of these topics, mental load, according to Muñoz (2012), is “**a construct of paramount importance**” for translation process research.

Significance: Practical

Knowing the difficulty level of a translation task is important in:

- translation pedagogy
- translation accreditation
- for the language industry



Significance: Practical

E.g.,

- According to the results of **expertise studies**, **deliberate practice** only occurs only under the following conditions, when (a) there is a well-defined task, (b) **the task is of appropriate difficulty for the individual**, (c) there is informative feedback, and (d) there are opportunities for repetition and the correction of errors (see Ericsson, 1996; Shreve, 2006).

Outline

- Significance
- **History**
- Concepts
- Research framework
- Measurement methods
- Research objects
- Some findings

History: In neighboring disciplines

- Since the 1920s, psychologists started to pay attention to workload and difficulty.
- Since the 1960s, mental workload has been an important concept in human factors and industrial psychology.



History: In Translation Studies

- **Analytical:** **Wilss** (1982) distinguished between four types of translation difficulty; **Nord** (2005) made similar distinctions: text-specific difficulties, translator-dependent difficulties, pragmatic difficulties, and technical difficulties.
- **Cognitive:** **Campbell**, S., 1999. A cognitive approach to source text difficulty in translation. *Target*, 11(1), pp.33-63.
- It suggests that difficulty can be interpreted in cognitive terms, with the support of models of working memory and of language comprehension and production.



History: In Translation Studies

- Empirical:
- Jensen (2009) employed three indicators to measure the translation difficulty level of some types of text.
- Liu and Chiu (2009) aimed at identifying indicators that may be used to predict source material difficulty for consecutive interpreting.
-

Outline

- Significance
- History
- **Concepts**
- Research framework
- Measurement methods
- Research objects
- Some findings

Concepts: Difficulty

According to Oxford English Dictionary,

(1) **Something** which is hard to carry out, deal with, or overcome; a **problem**.

E.g., What difficulties did you encounter?

(2) The **quality**, fact, or condition **of a thing** being hard to deal with or of presenting obstacles to progress or accomplishment.

E.g., How did you perceive the difficulty level of this task?

Concepts: Difficulty

Accordingly, difficulty in translation refers to

(1) translation problem

(2) the extent to which cognitive resources are consumed by a translation task for a translator to meet objective and subjective performance criteria (Sun, 2015)

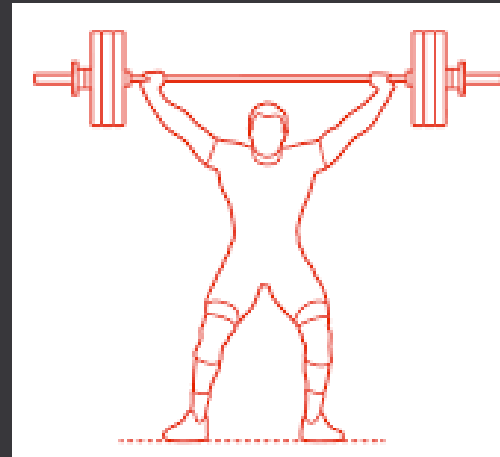
Concepts: Cognitive load & effort

- **Cognitive load**: The amount of mental effort demanded by a task
- **Cognitive effort**: The amount of mental effort put into a task by an individual (e.g., 60%? 90%? He did not do his best.)



Concepts: Cognitive load & effort

- Cognitive load is usually bigger than cognitive effort.
- In the case of translation (and interpreting), cognitive load often **equals** cognitive effort.



Concepts: Synonyms

- **Difficulty**, cognitive load, mental load, mental workload, cognitive workload, workload, cognitive effort, mental effort...
- **Difficulty** is broader in meaning than cognitive load or effort.




Outline

- Significance
- History
- Concepts
- **Research framework**
- Measurement methods
- Research objects
- Some findings

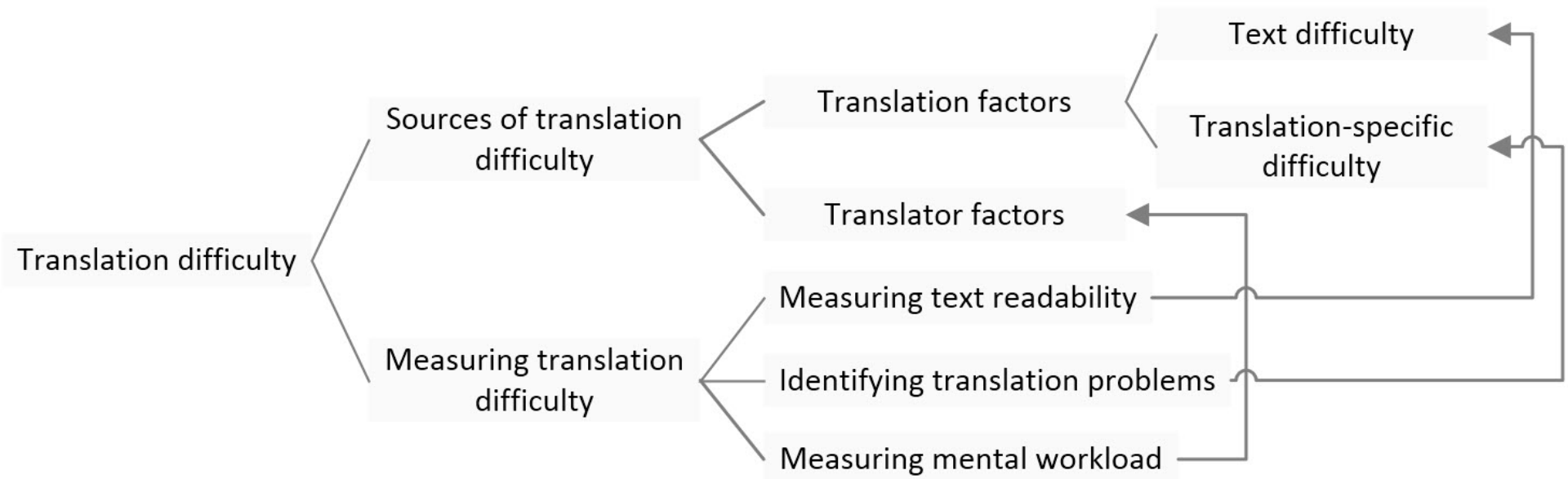
Framework: Research purposes

In translation difficulty research, two essential questions are:

- (1) what makes a text difficult to translate
- (2) how to **measure** and **predict** the difficulty degree of a translation task.

- **Measure:** “If you can’t measure it, you can’t improve it.” —Peter Drucker
 - **Predict:** Predict the difficulty level
- 
- Predict and model the behavior & performance
 - "Predictive Turn in Translation Studies: Review and Prospects" (Schaeffer, Nitzke, and Hansen-Schirra, 2020)

Framework: Translation difficulty research



Sources of translation difficulty

1. Translation factors include readability (or reading comprehension) problems and translation-specific (or reverbalization) problems
2. Translator factors concern translation **competence and expertise** (or “ability variables” such as intelligence, aptitude, cognitive style, and working memory capacity), and **affection** (or “affective variables” such as confidence, motivation, and anxiety).

Sources of translation difficulty

New 3. **task and environmental factors**, such as task criticality, task novelty (to the translator), time pressure, equipment used (e.g., CAT tools), physical and psychological environmental factors (ergonomics, UX).

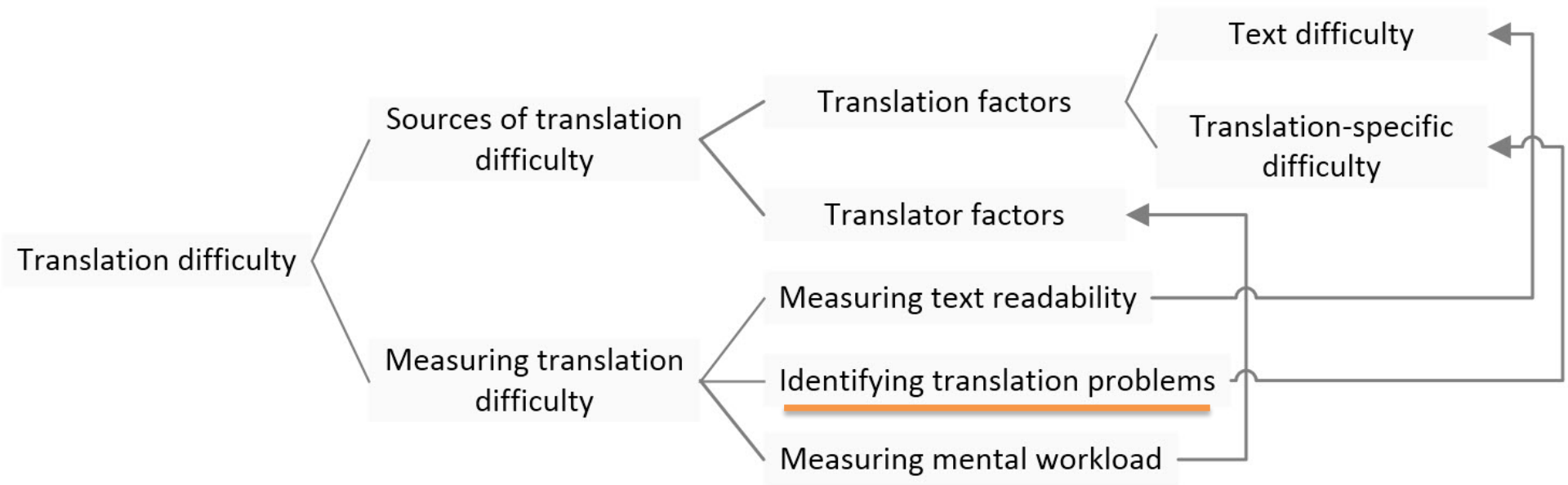
Text complexity

- Text Complexity (or readability) has been an important topic in reading research since the 1920s.
- Based on a survey, Gray & Leary (1935) compiled a list of **289 factors** that may be related to a book's readability.

Text complexity: Measurement

- To date, over 200 readability formulas (e.g., Flesch Reading Ease formula) have been published.
- The strongest indexes of readability are: **vocabulary difficulty and sentence length** (e.g., Chall & Dale 1995)

Framework: Translation difficulty research



Identifying translation problems

Methods:

- **Self-report**: Verbalization (e.g., thinking aloud) and reflective journal analysis
- **behavior analysais**: Pausing and revising behavior recorded by keylogging, eye-tracking or screen recording
- **Error analysis**

Identifying translation problems

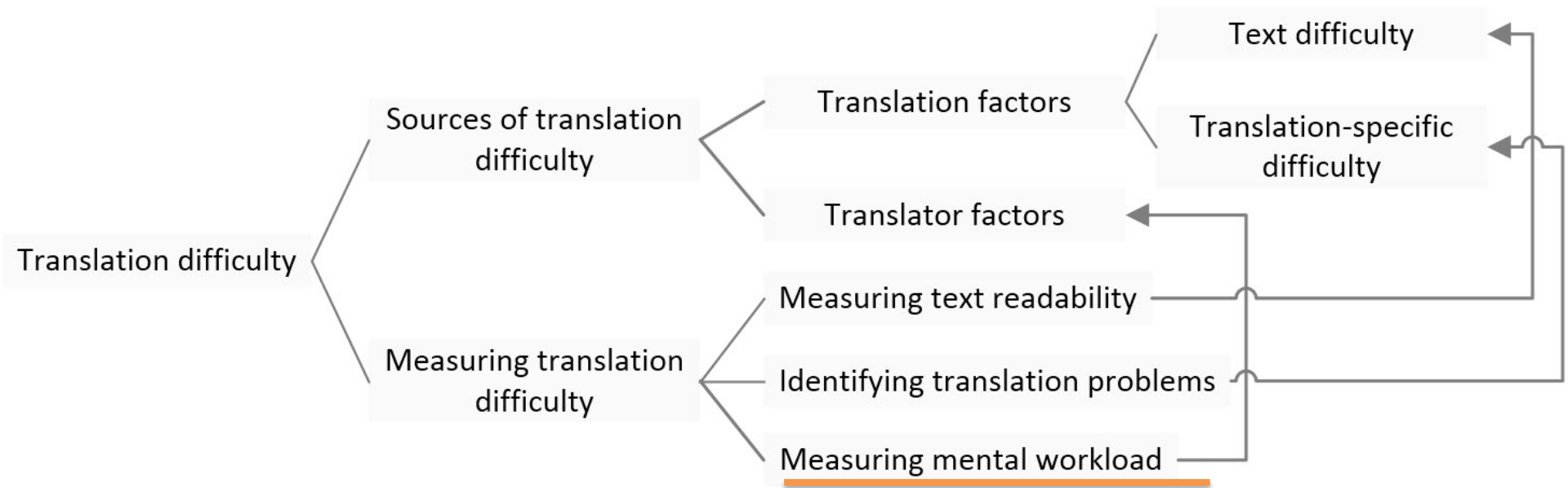
- Campbell and Hale (1999) identified several areas of difficulty in lexis and grammar, that is, words low in propositional content, complex noun phrases, abstractness, official terms, and passive verbs.

Identifying translation problems

Problem:

- We need to develop a process-oriented **typology of translation problems.**

Framework: Translation difficulty research



Measuring cognitive load or effort

- **Subjective measures** (e.g., rating scales, expert panel) (offline)
- Performance measures (e.g., grading translations, time-on-task) (offline & online)
- Physiological measures (e.g., EEG) (online)
- Behavioral measures (e.g., pausing recorded by **eye-tracking and keylogging**)

Subjective: NASA TLX rating scale

- **Mental Demand** How mentally demanding was the task?



- **Effort** How hard did you have to work to accomplish your level of performance?



- **Frustration** How insecure, discouraged, irritated, stressed, and annoyed were you?



- **Performance** How successful were you in accomplishing what you were asked to do?



Subjective: Pairwise comparison

- Pairwise comparison: any process of **comparing entities in pairs** to judge which of each entity is preferred, or has a greater amount of some quantitative property (Wikipedia).
- The **analytic hierarchy process** (AHP): a structured technique for organizing and analyzing complex decisions, based on mathematics and psychology. (Wikipedia)
- Can be used to judge, e.g., which text is more difficult, by a group of people.
- Software: Expert Choice, and others

- Pairwise comparison



- *The Social Network* (Movie: 2010)

Behavioral: eye-tracking & keylogging

- Eye tracking: the number of fixations, fixation durations, attentional switching, and scanpath similarity.
- Keylogging: Translog & Inputlog
- CRITT TPR-DB (see Carl et al, 2016)

To predict the difficulty level

- To find a valid and reliable method for **after-the-fact measurement** of difficulty for the individual
- To find an objective and automatic way to **predict** the workload independent of the individuals (like a readability formula)

To predict the difficulty level

- Use **machine learning** approaches to predict the quality, efficiency, and productivity of translation (see Schaeffer, Nitzke, and Hansen-Schirra, 2020).

Outline

- Significance
- History
- Concepts
- Research framework
- Measurement methods
- **Research objects**
- Some findings

Research objects: A typology

- Bottom Layer: Modes of translation
- Middle Layer:
 - translation subtasks
 - text factors
 - task constraints
 - translator factors
- Top Layer: cognitive effort, problem solving, translation strategies, translation styles, pauses, metacognition, uncertainty, emotion, mental representation, memory, and others.

Research Objects: Modes of translation

Translation difficulty research has been conducted in:

- written translation
- sight translation
- machine translation
- machine translation post-editing
- simultaneous or consecutive interpreting
- computer-aided translation (or interpreting)
- audiovisual translation
- ...

Research Objects: Middle Layer

- Four blocks: translation subtasks, text factors, task constraints, and translator factors.
- **Translation subtasks**: reading, drafting, seeking information, and revision.
- **Text factors**: translation directionality, text type (e.g., literary, specialized), and linguistic, textual and conceptual features of a text that a researcher can focus on, such as readability, cohesion, and potential problems (e.g., figurative language, metaphor, terms, speech disfluencies).

Research Objects: Middle Layer

- **Task constraints:** factors that researchers can manipulate, for example, access to reference resources and tools (e.g., electronic dictionaries), time pressure.
- **Translator factors:** translation competence and expertise levels, personality and other individual differences

Outline

- Significance
- History
- Concepts
- Research framework
- Measurement methods
- Research objects
- **Some findings**

Machine translation: difficulties

Google 学术搜索 machine translation difficulties

文章 找到约 653,000 条结果 (用时0.10秒)

时间不限
2021以来
2020以来
2017以来
自定义范围...

按相关性排序
按日期排序

不限语言
中文网页
简体中文网页

包括专利
 包含引用

创建快讯

Difficulties in establishing common ground in multiparty groups using machine translation
[N Yamashita](#), [R Inaba](#), [H Kuzuoka](#)... - Proceedings of the SIGCHI ..., 2009 - dl.acm.org
When people communicate in their native languages using **machine translation**, they face various problems in constructing common ground. This study investigates the **difficulties** of constructing common ground when multiparty groups (consisting of more than two language ...
☆ 被引用次数: 87 相关文章 所有 9 个版本

Six challenges for neural machine translation
[P Koehn](#), [R Knowles](#) - arXiv preprint arXiv:1706.03872, 2017 - arxiv.org
... but completely unrelated to the input, while the SMT output betrays its **difficulties** with coping ... from the learning curve experiments (Section 3.2), and used it to **translate** a collection of ... The key contribution of the attention model in neural **machine translation** (Bahdanau et al., 2015 ...
☆ 被引用次数: 671 相关文章 所有 5 个版本

[PDF] Word-sense disambiguation for machine translation
[D Vickrey](#), [L Biewald](#), [M Teyssier](#), [D Koller](#) - Proceedings of human ..., 2005 - aclweb.org
... They improve performance on the noun-phrase **translation** task, and show that they can use ... They present results which indicate that humans can accurately **translate** noun phrases without looking at ... paper, context can be very use- ful for a (sub-human-level) **machine** translator ...
☆ 被引用次数: 218 相关文章 所有 20 个版本

[PDF] Difficulties in processing malayalam verbs for statistical machine translation
[V Jayan](#), [VK Bhadran](#) - International Journal of Artificial Intelligence ..., 2015 - academia.edu
In this paper we discuss the **difficulties** in processing the Malayalam texts for Statistical

Written translation: Some findings

- Translation difficulty level and text readability were negatively and weakly related, which means that a text's readability only partially accounted for its translation difficulty level.
- A post-translation questionnaire survey showed that 77% of over 600 responses pointed to reverbalization in the target
- language as more difficult than source text comprehension. (Sun & Shreve, 2014)

Written translation: Some findings

- Liu, Zheng & Zhou (2019) explored the impact of text complexity on translators' subjective perception of translation difficulty and on their cognitive load.
- They found that (i) **text complexity** measured by readability, word frequency and non-literality was **in line with** the results received from informants' **subjective assessment** of translation difficulty; (ii) moderate and positive **correlations existed between** most items in the **self-assessments** and the indicator (fixation and saccade durations) obtained by the **eye-tracking** measurements

Post-editing: Some findings

Jia, Carl & Wang (2019) compared post-editing and from-scratch translation, and found that

1. post-editing was significantly faster than from-scratch translation for domain-specific texts;
2. post-editing significantly reduced the participants' cognitive effort.

...

Research: Bibliography

- Too many interesting studies to review in this talk.
- Intend to compile a **bibliography** on translation difficulty research
- You may **email** your works related to this topic to me (sunsanjun@bfsu.edu.cn), and I will **share** the bibliography with you in time.

Translation Difficulty Research

An Interesting and
Emerging Field

Thanks