ASSESSING THE QUALITY OF ENGLISH-TO-THAI MACHINE INTERPRETING:

TRANSLATING GOOGLE TRANSLATE
Abstract

This research aims to assess the quality and potential of Google Translate when used as an English-to-Thai Machine Interpreting (MI) tool. The paper explores MI’s distinct definitions and terminology as well as its relationship to other relevant technologies, i.e. Machine Translation (MT), Automatic Speech Recognition (ASR), and Text-to-Speech (TTS). Four five-minute English-language recordings from the European Commission’s Directorate General for Interpretation Speech Repository were selected as samples for the MI to interpret. Human evaluation was used to evaluate the Comprehensibility, Accuracy, and Usability of MI with five final-year interpreting students and six professional interpreters performing as judges. Results show that MI is still far from perfect with problems in every step of the MI process from ASR, MT, to TTS. Discussion and analysis further reveals how and why current MI fails in its processing of English into Thai, as well as presents contemplations on what should be done to enhance MI performance. The subtle decisions chosen, and mistakes made, by the MI tool when encountering uncertainty suggests that MI exhibit comparable error symptoms and display similar coping strategies to human interpreters. Thus, it is paramount that professional interpreters be involved the development and improvement of MI technology going forward.

Keywords: Machine Interpreting, Machine Translation, Automatic Speech Recognition, Text-to-Speech, Google Translate, English-to-Thai interpretation, comprehensibility, accuracy, coping strategies
1. Introduction

1.1 Background and Significance

In recent years corporations like Google and Microsoft have made great strides in Machine Interpreting. Google has incorporated voice recognition and text-to-speech into Google Translate, further advertising that their recently-released Pixel Buds peripheral can offer what they claim to be ‘real-time translation’. Google Translate is one of the few services that offers translation from English to Thai and vice versa.

Machine Interpretation is being advertised as the future of interlingua communication and may one day change how interpreters work. Winteringham (2010) suggests that if practitioners understand the potential of new technology and adapt accordingly it will benefit both themselves and the profession. However, the current tools available are deemed largely unsatisfactory and the translation quality remains the main concern (Gaspari, Almaghout, & Doherty, 2015; Zaretskaya, Pastor, & Seghiri, 2016). This is even more evident in less-common languages, such as Thai, where the corpus of reference translations is lacking and initiative for the research and development of translation technologies is inadequate (Wutiwiwatchai, Supnithi, & Kosawat, 2008).

Therefore, a study into the quality of Machine Interpreting and other components of the Machine Interpreting process of these tools is essential to the prospective use and improvement of the technology for practical interpreting. This research aims to assess the English to Thai Machine Interpreting quality of Google Translate.

1.2 Machine Translation vs Machine Interpreting

For the purposes of this study, it is important to make the distinction between ‘Machine Translation’ and ‘Machine Interpreting’.

*Machine Translation (MT)* is an automated process by which computer software is used to translate text input from a source language into a target language. This is the main method of automatic translation used in the past. It is also what is most commonly referred to when discussing automatic translation and the focus of the majority of the existing studies on this matter.

*Machine Interpreting (MI)* is the process of translating speech from the source language into target language speech. It is the integration of several different technologies into a system of automatic speech translation. The different technologies are namely Automatic Speech Recognition (ASR), Machine Translation, and Text-to-Speech (TTS). Artificial Intelligence (AI) and Machine Learning (ML) can also be integrated into such systems (Google, Microsoft) to improve the quality of translation over time. The MI process involves three basic steps:
1. Speech recognition software records a speaker’s voice and renders it into text.
2. The text is then translated into the target language using a MT program.
3. The translated text is automated into audio using a TTS program into the target language.

Nonetheless, the core of MI is still the MT component of the process and is an important part in assessing MI quality. This study focuses on evaluating the final rendered product as a means to assess the quality of MI while also considering the other factors and mechanisms involved in MI process that could impact its quality, namely the accuracy and speed of ASR, MT, and TTS rendition, and whether the intonation of synthesized speech influences the meaning of the interpretation.

2. Related Works

Most of the existing literature deals with MT (automated text-to-text translation) and there was very little research done on MI. However, as MI still relies heavily on MT, the methods used in past studies are relevant and can be applied this study.


The researcher reviewed approaches to MT quality assessment that have been studied in the works of Koehn (2004); Koehn et al. (2005); Koehn and Monz (2006), Bremin et al. (2010), Denkowski and Lavie (2010); Gilda (2017); Rowda and Pospelova (2016); Sabek (2014); Uszkoreit (2007); Zaretskaya et al. (2016); and Zhang (2016). Existing literature all agree on the key aspects of automated translation to be evaluated when making quality assessments, especially when using human evaluators, which can be summarised into three main parameters: Comprehensibility, Accuracy, and Usability.

Koehn’s works provided examples and guidelines for human evaluation. Kalyani, Kumud, Singh, Kumar, and Darbari (2014) and Seljan, Brkić, and Kućiš (2011) presented case studies of evaluating MT of English to Hindi and Croatian respectively.

While Wutiwiwatchai et al. (2008) show that there is considerable interest in speech-to-speech technology in Thailand for purposes of translation, there is still very little research on the quality of the translation tools currently available. The researcher was able to find only one study into the Quality of English-Thai MT. Lyons (2016) found that the translations were poor, with as much as 47.2% error words per translation.
2.1 Quality Assessment Methods

There are a variety of evaluation methods used to assess the quality of MT in the literature. The two main methods are Human Evaluation and Automatic Evaluation.

*Human Evaluation* is done by specialists such as language experts, linguists, or translators and interpreters.

*Automatic Evaluation* is the machine-automated process of MT evaluation that can be achieved using a variety of metrics and is often the preferred method because it is fast, cost-efficient, repeatable, and can be used on large sample sizes independent of language (Gilda, 2017; Koehn et al., 2005; Koehn & Monz, 2006).

While Automatic Evaluation is more convenient and preferable, the results tend to be unreliable (Koehn, 2004; Koponen, 2010). Quality scores can vary depending on which metric is used and lacks the in-depth analysis of Human Evaluation that this particular research desires.

However, using Human Evaluation is time-consuming, costly, subjective, and cannot be reused (Guzmán, Abdelali, Temnikova, Sajjad, & Vogel, 2015; Koehn & Monz, 2006). Koehn et al. (2005) and Doherty (2017) propose that the main challenge of using Human Evaluation is the establishment of clear definitions for adequacy and fluency so that evaluation is precise and consistent evaluations. Additionally, researchers must carefully consider the material and instructions provided to evaluators to ensure the “validity, reliability, and efficiency of the evaluation,” and “increase the value of the data” (Doherty, 2007 p. 11). Thus, it is necessary to fully understand the terminology that exists in previous works and to clearly define the definitions that will be used in this study.

2.2 Terminology

The three key parameters for quality assessment as surmised from the literature were Comprehensibility, Accuracy, and Usability. However, the terminology used in previous studies could potentially pose problems for evaluators due to their ambiguity. The terminology used in existing literature are as follows;

*Fluency* is the term commonly used in literature that refers to how well the translation is formed according to the grammatical and/or syntactic rules of the target language. It is simply a measure of how natural the translation sounds in the target language. However, fluency is also often used to indicate the clarity, comprehensibility, or understandability of the translation. This can be misleading as a fluent rendition may not always equate to a comprehensible one. Thus, the study used the term “Comprehensibility” to rate the level of which the rendition can be easily
understood, regardless of correctness to the source. This means the translation is judged without taking the source into account and ascertain if the interpretation is understandable on its own.

*Adequacy* is the term commonly used to refer to the accuracy or fidelity of the translation. It is the correctness of the translation when compared to the source. According to the Linguistic Data Consortium, it is defined as “how much of the meaning expressed in the gold-standard translation or source is also expressed in the target translation.” To avoid ambiguity in what can be deemed adequate, this research used the term “Accuracy” in its evaluation.

*Usability* is then determined by comparing the comprehensibility of the translation with its accuracy in conveying the message of the source. This is the judges’ decision of whether the MI rendition is suitable for practical use.

3. Methodology

3.1 MI Tool

This study assesses the quality of Google Translate as a free and readily available MI tool that offers English to Thai and Thai to English language support. Having been unable to obtain the Pixel Buds for testing, this research used the Google Translate mobile phone application on a Samsung Galaxy Note model SM-N935FZBOTHL.

3.2 Participants

Human evaluation was used to assess the quality of the MI tools tested. Recruited evaluators are commonly known as Judges and will be referred to as such from this point. The Judges were five final-year interpretation students from the Master of Arts Program in Interpretation at the Chalermprakiat Center of Translation and Interpretation, Chulalongkorn University and four professional interpreters.

3.3 Material

Four English Language speeches were used in this research, with each speech being approximately five minutes long for a total of 20 minutes of MI samples. The speeches were selected from the European Commission’s Directorate General for Interpretation Speech Repository. The Speech Repository is an e-learning tool developed by the DG Interpretation to help improve the quality of conference interpreter training. It is a vast collection of speeches, on varying subjects and with different styles, tailored specifically to the needs of student interpreters for practice in both consecutive and simultaneous.
Speeches in the Repository are categorised by Domain, Difficulty (basic, beginner, intermediate, advanced/test-type, and very advanced), use (simultaneous and consecutive), and type (conference, debate, hearing, interview, pedagogical material, and so on).

The M.A. in Interpretation at Chulalongkorn University ultimately prepares students for conference interpretation, an extremely challenging and skill-intensive field that could benefit from the use of machine assistance in the future. Determining that the Conference type would ideally be the ultimate benchmark for MI, it was selected for testing and evaluation. Speeches intended for simultaneous and consecutive practice were used. However, as the intended use proved to be irrelevant in this study, they were categorized by their difficulty level—two of which were selected for comparison. There were no conference speeches at the Basic level, hence the researcher selected those from the Beginner level and selected Advanced/Test-type speeches as the highest difficulty. The reasoning being that if MI is to replace human interpreters then it should be able to at least interpret the speeches used to test human interpreters. The researcher used the Speech Repository’s search tool and selected speeches that met the aforementioned criteria that were closest to five minutes in length. The four following speeches were selected:

1. **Advanced-level Conference Speech for Simultaneous Practice (A1)**
   
   “Work-life Balance Initiative”
   

2. **Advanced-level Conference Speech for Consecutive Practice (A2)**
   
   “EU’s Aid for Haiti”
   
   Description: Statement on the EU’s response to the Haiti earthquake disaster.

3. **Beginner-level Conference speech for Simultaneous Practice (B1)**
   
   “European Road Safety Day 2007”
   
   Description: Effects of Alcohol and Drugs on Road Safety – an introduction.

4. **Beginner-level Conference Speech for Consecutive Practice (B2)**
   
   “Beyond GDP: Measuring Progress, True Wealth, and the Well-being of Nations”
   
   Description: Appropriate indicators to measure growth.

With the limitations of the existing technology (which will be addressed in detail in the discussion) the sample speeches were broken into chunks of at least 10-30 seconds depending on
natural pauses in the speech or whenever the MI tool was forced to end a segment due to complications in its programming. This was in attempt make it practical for this study and theoretically enable the MI tool the ability to process and provide the best interpretation it can possibly achieve.

The segments were stitched together to form a complete rendition of the speech. MI renditions were then transcribed for the judges to evaluate alongside its audio playback. The original speeches were made available for judges to evaluate their accuracy after the comprehensibility evaluation is completed.

3.4 Definition and Scoring

The judges were asked to evaluate each of the interpretations using the following parameters, which were explained to the judges before the judging process began.

*Comprehensibility* was the term used to refer to comprehensibility and understanding. It was defined as the ‘ability of the interpretation to convey a clear and understandable message’. Judges were asked to rate “How well do you understand what the interpretation is saying, regardless of context?”

Comprehensibility was scored using a five-point scale that rates the rendition into Thai as having the following qualities:

- **5 points** for being **Fully comprehensible**
- **4 points** for **Good comprehensibility**
- **3 points** for **Average comprehensibility**
- **2 points** for **Low comprehensibility**
- **1 point** for being **Incomprehensible**

*Accuracy* was the term used to refer to adequacy and fidelity and was defined as the ‘amount of information or meaning from the source that the interpretation maintains’. Judges were provided the source speech and asked to rate “How much of the original message is maintained in the interpretation?”.

Accuracy was scored using a five-point scale that rates the rendition into the target language as having the following qualities:

- **5 points** for maintaining **All** the information/meaning
- **4 points** for maintaining **Most** of the information/meaning
3 points for maintaining an Average amount of information/meaning

2 points for maintaining Little information/meaning

1 point for maintaining None of the information/meaning

Usability was the term used to refer to whether the interpretation is fit for purpose and was defined as the ‘usability of the interpretation in place of human interpretation’. Judges were asked to rate “How usable is the interpretation, i.e. can it replace human interpretation?”.

Usability was scored using a five-point scale that rates the rendition into the target language as being usable at the following levels:

5 points for Full usability in place of human interpretation

4 points for High usability in place of human interpretation

3 points for Average usability in place of human interpretation

2 points for Low usability in place of human interpretation

1 point for being Unusable in place of human interpretation

Judges were asked to first evaluate the MI Comprehensibility by listening to the recorded MI. The judges were then asked to evaluate the MI Accuracy by watching the original clip while reviewing the provided transcript of the MI recording used for comprehensibility. Finally, the judges were asked to consider both the Comprehensibility and Accuracy to evaluate the MI Usability. This process was repeated for each sample clip. The researcher conducted post-evaluation discussions with the judges for an in-depth error analysis of the MI tool used regarding functionality, quality, as well as possible improvements that could be made to MI.

4. Results

Interpretation students rated the MI with an average score of 1.57 with a standard deviation of 0.67. Table 1 shows the scores for each parameter and sample speech.

<table>
<thead>
<tr>
<th>Interpretation Students (Chula)</th>
<th>Sample A1</th>
<th>Sample A2</th>
<th>Sample B1</th>
<th>Sample B2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehensibility</td>
<td>1.60</td>
<td>2.00</td>
<td>1.40</td>
<td>1.67</td>
</tr>
<tr>
<td>Accuracy</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>1.80</td>
</tr>
<tr>
<td>Usability</td>
<td>1.00</td>
<td>1.20</td>
<td>1.20</td>
<td>1.00</td>
</tr>
<tr>
<td>Average score</td>
<td>1.53</td>
<td>1.73</td>
<td>1.53</td>
<td>1.47</td>
</tr>
</tbody>
</table>

Table 1: Student scores
Professional interpreters rated the MI with an average score of 1.57 with a standard deviation of 0.70. Table 2 shows the scores for each parameter and sample speech.

<table>
<thead>
<tr>
<th>Professional Interpreters</th>
<th>Sample A1</th>
<th>Sample A2</th>
<th>Sample B1</th>
<th>Sample B2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehensibility</td>
<td>1.50</td>
<td>2.33</td>
<td>1.50</td>
<td>1.67</td>
</tr>
<tr>
<td>Accuracy</td>
<td>1.83</td>
<td>2.17</td>
<td>1.67</td>
<td>1.67</td>
</tr>
<tr>
<td>Usability</td>
<td>0.83</td>
<td>1.50</td>
<td>1.17</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Average score</strong></td>
<td><strong>1.45</strong></td>
<td><strong>2.00</strong></td>
<td><strong>1.44</strong></td>
<td><strong>1.44</strong></td>
</tr>
</tbody>
</table>

Table 2: Professional scores

The overall average score was 1.57 with a standard deviation of 0.62. Table 3 shows the scores for each parameter and sample speech.

<table>
<thead>
<tr>
<th>Overall</th>
<th>Sample A1</th>
<th>Sample A2</th>
<th>Sample B1</th>
<th>Sample B2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehensibility</td>
<td>1.55</td>
<td>2.18</td>
<td>1.45</td>
<td>1.64</td>
</tr>
<tr>
<td>Accuracy</td>
<td>1.90</td>
<td>2.09</td>
<td>1.82</td>
<td>1.73</td>
</tr>
<tr>
<td>Usability</td>
<td>0.91</td>
<td>1.36</td>
<td>1.18</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Average score</strong></td>
<td><strong>1.45</strong></td>
<td><strong>1.88</strong></td>
<td><strong>1.48</strong></td>
<td><strong>1.45</strong></td>
</tr>
</tbody>
</table>

Table 3: Overall scores

When considering scores by difficulty, the average score for Beginner level samples was 1.47 with students giving a 1.50 average score and professionals giving a 1.44 average score. Advanced level samples (Test-type) received a higher score of 1.67 with students giving a 1.63 average score and professionals giving a 1.73 average score.

5. Discussion

It is clear from the scores that MI in its current state cannot be used as a replacement for human interpreters, least of all in a conference setting which is more demanding than most other situations. The average scores never surpassed a 2.00 for Low Comprehensibility, Accuracy, and Usability meaning the MI renditions were incomprehensible, inaccurate, and unusable. Sample A2 scored highest among all samples possibly because it was a well-paced speech by a native speaker, unlike Sample A1 which was also given by a native speaker but at a faster pace. Samples B1 and B2 were given by non-native speakers with inconsistent pacing and pausing.

More than half the judges (52%) agreed that Comprehensibility and Accuracy was Low, with Little of the meaning being conveyed into the target language, while an additional 30% rated the renditions as being completely inaccurate and incomprehensible. Judges concurred that some
words and phrases were correctly translated, however without proper context and cohesiveness, none of them could be used and 86% of all MI renditions were deemed wholly Unusable.

To try to understand the low MI scores, the next section will attempt to analyse the shortcomings of MI by examining the different aspects of the MI process, what factors can affect quality, as well as the issues and challenges it faces.

5.1 Overall Issues and Challenges for Machine Interpreting

Many of the problems became evident in the early stages of testing the MI tool with the selected speeches. Some of these problems have been mentioned in related works on MT and have proven to carry over to MI to some extent.

5.1.1 Rendition Delay

The first issue this study faced was the MI delay in rendering the interpretation into speech. Despite the ability of MI to translate speech in real-time and even change the structure of the sentence as more context is presented, it is still unable to simultaneously render the translated text into speech. This also means that Simultaneous MI is impossible in its current state. For instance, it is possible to read the translated text in the mobile application as the speech is ongoing, however, the user is unable to listen to the interpretation until the MI considers the segment to be completed and stops recording for ASR. Moreover, the interpretation is only rendered into speech one segment at a time. Other speech segments cannot be interpreted in the background, meaning full speeches cannot be interpreted simultaneously by MI. This also indicates that current MI technology is more suitable for providing human interpreters or listeners in the audience with real-time translated text display than performing actual interpreting.

5.1.2 Automatic Speech Recognition

Further complications arose from the MI application’s inability to understand the speaker’s speech patterns and the distinguish the pauses in speech. The MI applications consider all pauses in speech longer than a certain amount to be the end of a sentence, regardless of the speaker’s intent or the completeness of the content of speech. In this study, the MI application would cut and end to the segment if there was between two to three seconds of dead air (2.67 seconds on average with high inconsistency). All samples contained cases when the speaker had not finished the sentence and the MI app started processing the incomplete input and rendered an incomprehensible translation. To further compound this problem, the MI tool lacks any control for the user to force the MI process to continue running through spells of silence and is therefore entirely reliant on the pre-programmed algorithms in deciding when segments end.
Speed was another important factor towards the quality of MI. The MI tool had difficulty following speakers due to the speed of delivery. For example, Sample B2 had a fast speaker who caused MI to stop responding and forced the researcher to restart the segment or break the segment into smaller chunks four separate occasions throughout the recording. However, this also happens when slow speakers frequently pause, hesitate, or use fillers. Too slow and the MI stalls, too fast and it pauses because it can’t catch up. Without human intervention, the MI is unable to modulate its processes to match the speed of the speech in the way human interpreters can when dealing different speakers.

Not only do those problems make it impossible to use the current MI tool for longer speeches, but it also affects the ability of the MI to create a comprehensive and cohesive interpretation. Without the ability to understand when a sentence ends the MI cannot adequately segment the speech into understandable pieces. Thus, an incomplete recording is rendered into text. When that text is translated and rendered into the target language, its comprehensibility and accuracy is inevitably lost.

Other technical problems involve ASR in situations of uncertainty, where speech is not picked up properly due to sound quality issues, an issue brought up by Mathias (2008); whereas issues mentioned by Wendt (2016) such as non-native accents, homophones, stress and intonation, and even names were observed to be particularly difficult for the ASR. One example is Sample B1 where the ASR failed to pick up simple phrases in the opening sentence, leading to a completely inaccurate interpretation.

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“The E.U. is not a market or a currency, those are instruments...”</td>
<td>“D u is not a market or a currency does our instruments...”</td>
<td>“D u ไม่ได้เป็นตลาดหรือสกุลเงินที่ใช้ในเครื่องมือของเรา…”</td>
</tr>
</tbody>
</table>

(D u is not a market or currency used in our instruments)

In Sample A2, once again the error occurs in the opening sentence and completely changes the meaning of the original speech.

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Our immediate focus is on working with the United Nations and the Haitian”</td>
<td>“Our naked focus is on working with the United Nations of a Haitian”</td>
<td>“เป้าหมายที่เปลือกษากรทำงานร่วมกับองค์กรสหประชาชาติในการเป็นผู้นำของ”</td>
</tr>
</tbody>
</table>
leadership...”  “leadership...”  ชาวเฮติ...”
(Our naked aim is working with the United Nations in leading the Haitians)

Despite the ASR correctly identifying the word “Haiti” earlier in Sample A2, it later mistakes it for “I-80” which sounds similar but completely loses the meaning of the source.

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“...we will <strong>of course</strong> travel to Haiti as soon as it is appropriate...”</td>
<td>“...we will <strong>have Coast</strong> travel I-80 as soon as it is appropriate...”</td>
<td>“…เราจะมีการเดินทางไปตามชายฝั่ง I-80 ทันทีที่เหมาะสม...”</td>
</tr>
</tbody>
</table>

This excerpt from Sample B2 is another misinterpretation example resulting from bad ASR.

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“We heard about civil society initiatives like local community associations...”</td>
<td>“we heard about secret societies initiatives like local communities associations...”</td>
<td>“เราจะได้ยินเกี่ยวกับการเริ่มต้นของสังคมลับเช่นสมาคมชุมชนในท้องถิ่น...”</td>
</tr>
</tbody>
</table>

(We heard about the initiatives of secret societies like local community associations...)

5.1.3 Machine Translation

ASR is intrinsically linked to this step and the failure of the former produced problems in the latter. Okpor (2014) and Wendt (2016) argue that problems lie within the MT programs themselves. Existing MT algorithms designed to translate written texts have trouble handling disfluencies in human speech such as missing punctuations and conversational filler. Thus, the MI’s rendition of speech-to-text posed a significant hurdle as the entirety of the text rendered from the source speech contained no punctuation.

When given the ASR transcripts, the main complaint from judges is that the speech-to-text and the subsequent MT is produced without any punctuation or separation of ideas, often stringing
multiple sentences together, losing any cohesiveness. While the ASR did well to avoid filler, its lack of punctuation meant that it was unsurprising for there to be problems in MT. Sentences were garbled and meaning was lost when several ideas were expressed without any clear distinction between them.

Sample B1 is a prime example, if we ignore the interpretation errors and focus on punctuation, the researcher identified, and indicated in parenthesis, that there are at least six separate ideas in the original passage. However, as the ASR is incapable of rendering punctuation, the entirety of it is translated and rendered into Thai without any pauses in the audio rendition from TTS. This, according to the judges, was very confusing to listen to and nearly impossible to make sense of when considering the numerous other errors.

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“(1) Before we start, <em>again a reminder</em>, the Red Cross is saying have still capacities for you to refresh your first aid skills. (2) You just go there, you don’t need to be inscribed, <em>(3)</em> if you <em>are subscribed</em> that’s good, <em>(4)</em> if not you just go there, and they’ll be happy to teach you one or two <em>drinks</em>. <em>(5)</em> Take advantage of it, it’s really something which I think is <em>this</em> useful <em>(6)</em> and of course, free of charge.”</td>
<td>“[..] Begin a reminder the Red Cross is saying they have still capacities for you to refresh your first aid skills you just go there you don’t need to be inscribed if you unsubscribe is good if not you just go there and I’ll be happy to teach you one or two drinks take advantage of it it’s really something which I think is this useful and of course free of charge”</td>
<td>“เริ่มเต้นการเตือนสติกาชาดบอกว่าพวกเขายังคงมีขีดความสามารถใน การฟื้นฟูทักษะการปฐมพยาบาลของคุณที่คุณเพิ่งไปที่นั่นคุณไม่จำเป็นต้องถูกจารึกไว้ถ้าคุณยกเลิกการเป็นคนดีถ้าคุณไม่ไปที่นั่นและฉันยินดีที่จะสอนคุณหนึ่งหรือสองเครื่องดื่มใช้ประโยชน์จากมันนับเป็นจริงสิ่งที่ฉันคิดว่าเป็นประโยชน์และแน่นอนเสียที phí”</td>
</tr>
</tbody>
</table>
However, Ballabh and Jaiswal (2015) note that the different approaches to MT, e.g. rules-based, corpus-based, or statistical models, and the rest can result in vastly inconsistent renditions of the same source material. Even when the speech is correctly recorded into text, some of the translation choices still did not make much sense. An example of this is in Sample B2 when the MI was able to record the commonly used “without further ado”, but failed to translate it correctly into Thai. It instead elected to transliterate “Ado” leaving “without further” to be transcoded into “โดยไม่ต้องเพิ่มเติม” (doj mài tōng pʰjâːm tɕːm) “without adding”, showing that it did not grasp the overall meaning of the source.

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“…and we start, without further ado…”</td>
<td>“…and so we start without further Ado…”</td>
<td>“…และเพื่อให้เริ่มต้นโดยไม่ต้องเพิ่มเติม Ado…”</td>
</tr>
</tbody>
</table>

In the same sample, the MI was able to correctly record the phrase “novice driver” but chose a questionable Thai translation that completely misinterprets the original speech.

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“…recommends zero-level for novice drivers…”</td>
<td>“…recommend zero level for novice drivers…”</td>
<td>“…แนะนำให้เป็นศูนย์ระดับสำหรับไตรตรการสามเณร…”</td>
</tr>
</tbody>
</table>

A further example in this sample clip is the correct recording of the phrase “24/7” for “twenty-four hours a day, seven days a week”. However, as it was transferred into text as the numbers 24/7, the Thai rendition vocalised this as “jìːsipːjíːs uːan tɕːtʰ” literally “twenty-four parts of seven” meaning ‘twenty-four sevenths’ instead, pointing to a miscomprehension of the context in the original speech.

Any of the issues mentioned above can alter the meaning of the message and, confirms what Anumanchipalli, Oliveira, and Black (2012) point out, that it can result in a loss of not only information and meaning but also the intent of the original source speech.
5.1.4 Text-to-Speech

In the judges’ view the audio rendition of the MI was simply unnatural. Apart from the failings of the ASR as mentioned above, the Thai language posed its own problems with the limited punctuation in the target language. Therefore, the rendered speech predictably inherits the same problems. Sentences tended to pause at awkward points, such as mid-sentence or even mid-word, which often caused confusion. There were a total of ... inappropriate pauses, averaging ... inappropriate pauses per sample speech.

For example, in Sample A1, the Thai phrase “การดูแลระยะยาว” (khrā? jā? jāw) “long-term care” was rendered as “การดูแล [pause] ระยะ” (khrā? [pause] jā? jāw); whereas “ระยะ” (jā? jā?) “term” is one word.

Sample A1 contains the phrase “เด็ก” (dek) “children”. The ๆ symbol or ‘mai ya-mok’ is an iteration mark representing a doubled word. However, instead of repeating the preceding word, this phrase was rendered into “เด็ก มายมะก” (dek maij ja:mok) “child ‘iteration mark’”.

Sample A2 sees a rendition of the word “นิ” (naj) “in” as “ไล” [pause] นิ” (laj [pause] nê) inexplicably separating the vowel ใ (laj) from the consonant น (nê).

Another example is in Sample B1 where the Thai phrase “เครื่องดื่มแอลกอฮอล์” (khrā?mok jā: œjá:kõː) “alcoholic beverages” was rendered as “เครื่อง [pause] ดื่ม แอลกอฮอล์” (khrā? [pause] dêj mok jā: œjá:kõː); whereas “เครื่องดื่ม” (khrā?mok jā:) “beverage” is one word. The point of separation is illogical because เครื่อง is a compound word formed from ‘เครื่อง’ (khrā?) ‘thing’ and ‘ดื่ม’ (dêj) ‘drink’, literally ‘thing [to] drink’ i.e. ‘beverage’, and if a mistake were to be made it would likely be separating these two words. Instead, the MI chose to separate เครื่อง into ‘เครื่อง’ (khrā?) and a lone consonant จ (jā) neither of which have any meaning.

Sample B2 shows us another inappropriate pause when it renders the word “เรา” (raw) “us” as “เรา [pause] อา” (re: [pause] jàː) mistaking the vowel ใ (laj) encompassing the consonant ำ (laj) for two vowels ใ (laj) and ำ (laj). Once again, the point of separation is confusing because while ‘เรา’ (re:) is a proper sound, albeit with no meaning, the vowel ำ cannot be used on its own without a consonant. This underscores the gap in machine comprehension of the Thai language.

Judges also commented that while a more natural voice is preferable, it is not the most urgent issue as the monotone did not greatly affect the comprehensibility of the rendition.

5.2 Coping Strategies

Possibly the most interesting result of the study is in seeing how the MI algorithm handles the problems outlined in the previous section (5.1). Even professional human interpreters are
prone to misinterpreting words, segments, or entire concepts. While it may not be possible to directly compare the root cause of the issues and challenges human interpreters are faced with to those encountered in MI, the manifestations of the symptoms and strategies used to combat them are surprisingly similar. One can draw parallels between the issues and challenges mentioned in 5.1 with the issues and challenges that human interpreters confront on the job.

The researcher was able to observe this because the Google Translate MI process is displayed in real-time, giving a glimpse into how MI copes with challenges. One can witness the ASR rendering speech-to-text and the MT translating that text as it happens. The recorded words and translated text updates as new information is obtained, altering the word choice and sentence structure to adapt the MI tool’s new understanding of the source speech. Problem symptoms and coping tactics could, therefore, also be seen in real-time as they occur.

At times when the speech was either too fast or contained words/names that the ASR could not clearly detect, Google Translate was observed to wait for more of the speech before proceeding. It would then perform the following actions to cope with the situation:

a.) End the current process and render what was recorded before this point, resulting in an incomplete interpretation.

b.) Record the closest estimation or approximation of that word/phrase into text for translation. The selected extracts from all Samples (A1, A2, B1, and B2), as shown below respectively, exhibit this error. These include instances of the MI having to guess what the speaker is saying. The first example for Sample A1 and the example for Sample B1 are cases in which the MI employs the transliteration of words it was unable to translate due to faults in word-detection.

<table>
<thead>
<tr>
<th>Sample A1:</th>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I think Europe has a role, in protecting and strengthening the rights of all our people...”</td>
<td>“I think you have a role in protecting and strengthening the rights of hola people...”</td>
<td>“ฉันคิดว่าคุณมีบทบาท ปกป้องและเสริมสร้างสิทธิของคน hola...” (I think you have a role in protecting and strengthening the rights of hola people...)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample A2:</th>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“…we see around us, a so-called counter-revolution”</td>
<td>“…we see around us so cool cultural counter-revolution”</td>
<td>“…เราเห็นรอบตัวเราวาดังนั้นการปฏิวัติทางวัฒนธรรมแบบย้อนยุคใน…”</td>
<td></td>
</tr>
</tbody>
</table>
emerge…”

Magic…”

เวทมนตร์…”

(...we see around us therefore the retro revolution of culture in magic...)

Sample A2:

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“It called for, <em>in due course</em>, and after post-emergency needs have been fully assessed...”</td>
<td>“It called for <em>intercourse</em> after post emergency needs to be fully assessed...”</td>
<td>“มันเรียกร้องให้มีเพศสัมพันธ์หลังจากฉุกเฉินโพสต์จ้าเป็นดังได้รับการประเมินอย่างเต็มที่...”</td>
</tr>
</tbody>
</table>

Sample B1:

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“...as we heard this morning, and <em>saw in</em> this very useful report card on member states...”</td>
<td>“...as we heard this morning and <em>Soarin</em> report card on member states...”</td>
<td>“...ในขณะที่เราได้ยินเมื่อเช้านี้และ <em>Soarin</em> บัตรรายงานเกี่ยวกับรัฐสมาชิก...”</td>
</tr>
</tbody>
</table>

Sample B2:

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“...not just <em>European Parliament</em>, to associate themselves to <em>this effort</em>...”</td>
<td>“...<em>or just you</em> parliament to associate themselves to <em>decipher</em>...”</td>
<td>“...หรือเฉพาะรัฐสภาที่จะเชื่อมโยงตัวเอง <em>แปลรหัส</em>...”</td>
</tr>
</tbody>
</table>
c.) Skip the unintelligible portion and completes the current sentence while omitting the skipped portion. For example, Sample A1 omits the entire portion in brackets.

<table>
<thead>
<tr>
<th>Original Speech</th>
<th>ASR Transcript</th>
<th>Thai Translation</th>
</tr>
</thead>
</table>
| “...the pillar of social rights and the principles, [that we believe are important to help member states to reform], and together converge upwards...” | “...the pillar of social rights and the principles [...] and together converge up words...” | “...เสาหลักของสิทธิและหลักการทางสังคมรวมตัวกัน...”

(...pillar of social rights and principles combined... [the ASR mistake ‘up words’ was omitted in the final rendition])

The examples above resemble some of the coping strategies used by professional interpreters. These are mentioned in Bartłomiejczyk (2006), Li (2015) and Setton and Dawrant (2016), as approximation, substitution, omission, transcoding, and reconstruction; which can result in the same misinterpretations evidenced in MI. Nevertheless, these solutions are not ideal, especially in reconstructing segments and omissions which rely on context and should always be selective and conscious, based on the interpreter’s judgement (Setton & Dawrant, 2016), something MI does not possess.

6. Conclusion and Recommendations

In its current state, Machine Interpreting (MI) is completely unusable for English to Thai interpretation and cannot replace human interpreters. Judges agree that the Comprehensibility, Accuracy, and Usability of all the MI samples used in the study were substandard, scoring a combined average of only 1.57 out of a possible 5.00, i.e. less than 32%.

Contributing factors are present throughout the entire MI process, but the main problem this research identified is the very first step: Automatic Speech Recognition (ASR). The inclusion of more robust recording options, such as a better algorithm for the ASR of longer speech, or manual control over the length of speech recording will be a step in the right direction. Better ASR is required, especially when dealing with non-native speakers’ accents, which demands a focus on more than just word detection. The ASR’s failure to function adequately causes a cascade of problems for all the other steps in the MI process. Greater understanding of the context of the source speech would greatly improve ASR and allow for better speech-to-text which, at the moment, crucially lacks punctuation and separation of ideas, which judges remarked as being the second largest flaw in the MI process.
Improvements to the Machine Translation (MT) and more natural Thai language text-to-speech (TTS) were recommended by judges, but it was noted that they are not the most urgent issues that needed to be addressed. Nevertheless, judges commented that there is definitely potential for MI to improve, especially with the use of Artificial Intelligence (AI) and Machine Learning (ML) in modern MI tools to mainly assist, or even potentially replace, human interpreters in the future. It was evidenced that MI demonstrates many of the error symptoms and coping strategies that can be noticed in human interpreters. Therefore, it is highly recommended that human interpreters play a key part in MI-technology development going forward.

Further research into the quality of MI is necessary for improving the technology. Although time-consuming, testing more varied sample types with a greater volume of samples and larger pool of evaluators will increase our understanding of the potentials and pitfalls of practical MI for English to Thai. While this research relied on established evaluation methods in the field of MT, it could be beneficial for future studies to investigate the practicality of developing novel evaluation processes and error analysis systems that are better suited for MI assessment. Moreover, this research focused on only one MI tool. Other MI tools should be assessed for example Microsoft Translator, the second most popular free translation tool with MI potential, which is also the algorithm used by many third-party translators.

This research recommends that software developers work closely with professional interpreters and translators to identify ways to better MI and provide a more natural experience. The first step in this direction would be to improve the quality and accuracy of ASR. In particular, it is essential to have the ability for MI to process longer speech segments or have the option to force the recording to continue until the segment is deemed complete by either the AI algorithm or a human user. This will all help to improve the ability of MI at the ASR stage to better understand the context of the speech. As Goldman-Eisler (1972, cited in Setton, 1999) highlights, in simultaneous interpreting, comprehension of the speech is what allows for accurate segmentation to be achieved. It is crucial aspect to correct because the most common criticism against MI from all the judges is its inability to analyse the context of the speech segment that is being translated.
References


Appendix A

Sample A1 Transcript

"Work-life Balance Initiative"

The EU is not a market or a currency, those are instruments. It's not just about consumers who consume, it's about citizens, real people, who try to make a living, getting their kids through school, struggling for a better future. I think Europe have has a role in protecting and strengthening the rights of all our people, and are we are not a day too late to do so. In the wake of the White Paper on the Future of Europe, the first reflection paper in the social dimension, has now been published, setting out scenarios on what choices we can make together. And we are also here today to speak about the pillar of social rights and the principles that we believe are important to help member states to reform and together converge up words.

Valdis Dombrovskis will give a detailed presentation on both the reflection paper in the pillar and Marianne Thyssen will conclude on behalf of the commission. For now let me focus on one of the forward to put the principles of the pillar into action. It’s a joint proposal Marianne Thyssen and Vera Jourová that I wish to attach great importance to, and one of which I'm very proud, frankly. The commission today adopted an ambitious new initiative to support Work-life Balance for both parents and carers. With this package of legislative and non-legislative action, we want to increase women’s participation and representation in the labour market and ensure that they’re not oblige, any more, to choose between a family or a professional career.

And this is not just women, it's about men as well. Existing policies have not brought equal opportunities that allow fathers and mothers to work and care together for the welfare of the children, elders, and society at large. Today the employment gap between men and women is more than 11% and in full-time equivalent it goes above 18% because part-time workers are mainly women, and when families have young children the gap widens hugely. Women also spend five times more time than men on care and household activities.

This explains partly the pay gap of 16% and the pension gap of 40% which increases women's exposure to poverty and economic dependence. The main drivers are the problem relate to parenthood; the unequal share of care responsibilities, to shortage of childcare and long-term care facilities, and financial disincentives in tax benefit systems for the second earner.
The economic loss due to the gender employment gap is estimated to be around 370 billion euro per year. In the context of demographic aging, we cannot afford leaving this huge talent untapped. This is a social and economic imperative and of course, above all, a moral imperative.

Our proposal, once adopted, will repeal the parental leave directive to replace it by a directive for work-life balance for parents and carers. We propose maintaining the existing four months parental leave entitlement for each parent but paid at sick-pay level and non-transferable. We introduce a paternity leave of ten days and carer leave five days per year and per dependent relative, also paid at sick-pay level. And we also extend and modernise the use of flexible work arrangements. At the same time, we will promote the provision of affordable and quality formal care services across the union including by optimizing the use of available EU financial support.

We will ensure better implementation and effective enforcement of the protection against dismissal. Provide guidance the member states and monitor their work-life balance policies development. Honourable members, we must stop wasting all that female talent and also give fathers a real chance to be with their children as well. We are in times, when we see around us a so-called cultural counter-revolution emerging, throwing us back into the past often at the detriment of women. I’m more than convinced that we need this package now. I do not want my daughters to be in the same situation as my grandmothers were.

Europe is not telling anyone how to live. There’s not one right balance but I do not accept that I don’t have to fight the battle again, that their grandmothers won. Work-life balance is about a real choice fathers and mothers now to organize our lives. This is our message: it’s great to be a mum, it’s cool to be a dad, and it is fantastic to be able to combine both roles with a carer, thank you very much.

D u is not a market or a currency does our instruments it's not just about consumers who can see him it's about citizens real people or try to make a living getting the kids to school struggling for a better future I think you have has a role in protecting and strengthening the rights of hola people are we are not a day too late to do so the wake of the white paper in the future of Europe the first reflection paper the social I mentioned is not be published setting out scenarios on what choices we can make together and we are also here today to speak about the pillar of social rights and the principles and together converge up words
reflection paper in the pillow and for now let me focus on one of the initiatives will bring forward to put the principles of the pill interaction it's a Droid purposive myint S & V are you over that I wish to attach great and I'm very proud Frankie the commission today doctors in ambitious New Balance for but parents and carols
where this package is legislative and non-legislative action you want to increase women's participation and representation in the labor market and it sure they're not oblige any more to choose between a family or professional career and this is not just about as well
existing policies if not brought equal opportunities that allow fathers and mothers to

D ไม่ได้เป็นตลาดหรือสกุลเงินที่ใช้ในเครื่องมือของเรามิ่งายตัวผู้บริโภคที่สามารถมองเห็นเขาได้เกี่ยวกับคนจริงของประชาชนหรือพยายามหาลี้ของชีพให้เด็กๆไปโรงเรียนที่เด็กจะอัมพาทตามที่ยังไม่ได้คิดว่าคุณมีบทบาท ปกป้องและเสริมสร้างสิทธิของคน hola เราไม่ได้เป็นวันที่สายเกินไปที่จะทำแข่งกับการปลุกกระดาษสีขาวในอนาคตของยุโรปกระดาษสะท้อนครั้งแรกที่สังคมที่กล่าวถึงจะไม่ถูกเผยแพร่การตั้งค่าสถานการณ์สิ่งที่ตัวเลือกที่สามารถ ร่วมกัน และเรายังยุติที่จะปัจจุบันเพื่อชูให้แสงหลักของสิทธิและหลักการทางสังคมรวมด้วยกัน

กระแสจะท้องในเหมาะสมและตลอดที่มีปลื้มผู้นำนั้นไปข้างหน้าหนึ่งในการวางหลักการของการกระทําสําหรับตลอดที่มีปลื้มผู้นำไปที่หนึ่งในความคิดริเริ่มที่จะเปลี่ยนการวางหลักการของปฏิสัมพันธ์ยานพาหนะเป็น Droid celinting myint S & V คุณมากกว่าที่อันค้องการที่จะแนบที่ดีและขันภูมิใจมากพร้อมกับคณะกรรมการในวันนี้แทนที่ ความสมดุลใหม่ที่มีความหมายจะยืนสำหรับแพ้มะและน้าร็อง

กรณีที่แตกก็จะนี้เป็นการกระทําที่ถูกกฎหมายและไม่ใช้กฎหมายที่คุณต้องการเพื่อการมีสวนรวมและ
การเป็นตัวแทนของชีพในตลาดแรงงานและแน่ใจว่าพวกเขาไม่จำเป็นต้องเลือกระหว่างอาชีพของครอบครัวหรืออาชีพมากกว่านี้

นโยบายที่มีอยู่ก็ไม่ได้นำไปสู่ที่ทางบุกันที่จะช่วยให้บุคคลและมาจากทางงานร่วมกันและเพื่อกันและ
work and cat together for the welfare of the children elders and Society at large today the employment gap between men and women is more than 11% and then full time equivalent it goes above 18% because part-time workers a menu women and when families have young children the Gap why didn't you g women also spend five times more than men on care and household activities

This explains about the pay gap of 16% and the pension gap of 40% which increases women's exposure to poverty and economic dependence the main drivers are the problem related Parenthood the unequal sharing of responsibilities to shortage of childcare and long-term care facilities and Financial disincentives in tax benefit systems for the second owner

The economic loss due to the gender employment gap is estimated to be around 370 billion euro per year in the context of demographic aging we cannot afford leaving this huge Talent untapped this is a social and economic imperative and of course above all the moral imperative

A proposal wants a ducted will repeal the parental leave directive to replace it buy it directly from work work life balance for parents and carrots we propose maintaining the existing 4 months parental leave entitlement for each parent but pay that sick pay level and non-transferable

ยังเพื่อสวัสดิการของการเด็กผู้สูงอายุและสังคมในปัจจุบันช่องว่างภาวะการจ้างงานระหว่างชายและหญิงมีมากกว่า 11% และและยังเท่าเดิมเวลาแล้วจะเกินกว่า 18 ปี เพราะคนทำงานนอกเวลาเป็นผู้หญิงและเมื่อครอบครัวมีเด็กเล็ก Gap ทำให้คุณมา หญิงผู้หญิงจึงใช้เวลาทำครัวมากกว่าผู้ชายในการดูแลและกิจกรรมที่ใช้ในครัวเรือน

ช่องอิทธิพลของช่องว่างด้านค่าจ้างที่ 16% และช่องว่างด้านบ้านจ่ายที่ 40% ซึ่งจะเพิ่มความเสี่ยงต่อการขาดแคลนของผู้หญิงและการพึ่งพาอาศัยกันทางเศรษฐกิจโดยตรงล่าสุดคือปัญหายังเกิดขึ้นความเป็นปัญหาในการแบ่งปันความรับผิดชอบที่ไม่เท่าเทียมกันการขาดแคลนการดูแลเด็กและการดูแลระยะยาว ลดหย่อนภาษีในระบบประโยชน์ทางภาษีสำหรับเจ้าของคนที่สอง

การสูญเสียทางเศรษฐกิจอั้นเนื่องมาจากช่องว่างภาวะการจ้างงานของเพศหญิงอาจจะอยู่ที่ประมาณ 370 พันล้านยูโรต่อได้ปีของความชราภาพทางประชากรที่เราไม่สามารถจะที่ก็ความสามารถพิเศษนี้ให้ได้ก็คือความจำเป็นทางสังคมและเศรษฐกิจและเห็นว่าความจำเป็นทางศีลธรรมทั้งหมด

ข้อเสนอแนะว่าต้องการยกเลิกสัญญายกฟรีให้เฉพาะจากผู้ปกครองเพื่อทดแทนการทำงานชื่อโดยตรงจากการทำงานในชีวิตการทำงานสมดุลสำหรับพ่อแม่และครอบครัวที่เราเสนอให้การปรับใช้ในการเปลี่ยนผู้ปกครอง 4 เดือนที่มีอยู่สำหรับผู้ปกครองต่อราย แต่จ่ายเงินค่าจ้างนั้นและไม่สามารถโอนได้
We introduce a paternity leave in 10 days in Carolee 5 days per year and per dependent relative also paid at sick pay level to use a flexible work Arrangements the same time we will promote the provision of affordable and quality formal care services across the union including by optimizing the use of available EU financial support.

Will it sure if that's an implementation and effective enforcement of a protection against this missile Guidance the member states and monitor their work-life balance policies development members we must stop wasting all that female talent and also get father’s a real chance to be with the children as well we are in times when we see around us so cool cultural counter-revolution in Magic throwing us back into the past often at the detriment of women and more than convinced that we need this package now I do not want my daughters to be in the same situation as my grandmother’s well you’re not telling anyone how to live there’s not one right balance but I do not accept that I don’t have to fight the bathroom again that their grandmothers one where can I find answers about a real Choice fathers and mothers now to organize our lives this is a message straight to be your mom it’s cool to be a dad and it is fantastic to be able to combine both roles with a Cara thank you very much.

จะตรวจสอบว่าเป็นการปฏิบัติและการบังคับใช้อย่างมีประสิทธิภาพในการป้องกันข้าpiesบนนี้แนะนำวัตถุสมัยและตรวจสอบการพัฒนานโยบายด้านสมดุลชีวิตการทำงานของพวกเขาเราต้องเห็นการคิดเรื่อง yawเสียงความสามารถทางเพศหญิงที่มีผลกระทบและทำให้ข้อความของอากาศจริงที่จะอยู่กับเด็ก ๆ เราอยู่ในช่วงเวลาที่เราเห็นรอบตัวเราดังนี้การปฏิวัติทางวัฒนธรรมแบบย้อนยุคในเวทมนตร์ที่ขว้างปาเราล้มเข้าในไปยังนิติมักเป็นการคิดเรื่องของผู้หญิงและมากกว่าที่เราเชื่อว่าเราต้องการคิดนั้นนั้นไม่ต้องการให้ลูกสาวของฉันเข้ามา สถานการณ์เช่นเดียวกับคุณยายของฉันตีคุณไม่ได้บอกทุกคนว่าจะมีชีวิตอยู่ที่นี่มีความสมดุลไม่เท่าไร แต่ลิ่วอยู่รับว่าฉันไม่ต้องผู้ที่อยู่ด้านอีกครั้งว่าคุณยายของเธอสิ่งต้องการคิดคำตอบเกี่ยวกับบรรพบุรุษและคุณยายที่แท้จริงได้ในขณะนี้เพื่อจัดระเบียบ ชีวิตนี้เป็นข้อความตรงไปที่จะเป็นฉันของคุณมันจะเป็นหน่อยและเป็นที่ยอดเยี่ยมเพียงพอให้สามารถที่จะรวมทั้งปายนายก Cilla ขอบคุณมาก.
Appendix B

Sample A2 Transcript

“EU’s Aid for Haiti”

Our immediate focus is on working with the United Nations and the Haitian leadership to ease the suffering of the Haitian people. Our commitment to rebuilding Haiti is long-term. Many European citizens have also lost their lives and around 1,000 are still unaccounted for. The EU has responded quickly, not seeking headlines but with the total focus on getting help to the people in need.

On advice from the United Nations be resisted the urge to travel to Haiti immediately, that would only have diverted attention and scarce resources away from the relief effort. We will of course travel to Haiti as soon as it is appropriate, and I’ve agreed with commissioner de Groot that he will travel to Haiti this week to extend the EU’s condolences and underline our commitment to the people. He will also take the opportunity to review our aid efforts so far and discuss with the United Nations and the people on the ground, the most pressing needs for the weeks and months to come.

In the meantime, we continue to work flat-out on all tracks: humanitarian, political, and security. Over the past days I have been in constant contact with Secretary of State Clinton, with the United Nations leadership, with EU foreign ministers, and with Canada which is leading the Friends of Haiti group, all this for the view to insure an effective and coordinated International response.

I will travel to the United States this week to follow up on this and other issues with the United States Administration and with the UN secretary-general and others in New York. The UN has asked for Urgent financial aid, 575 million U.S. dollars as well as logistical assistance to transport humanitarian aid yesterday the United Nation secretary-general ask for police and military reinforcements of the UN Peacekeeping Mission as well.

To help mobilize on coordinate our response, I asked the Spanish EU presidency to call an extraordinary Foreign Affairs council yesterday. For the first time since the Lisbon treaty entered into effect, we’re now pulling together the efforts of the commission and the council Secretariat, and the member states in a comprehensive approach under my overall coordination. This is ground-breaking. We had a productive council meeting yesterday. Everyone agreed on the need for rapid response and close coordination with the United Nations.
In terms of financial assistance, the council reach the following conclusion. It welcomes the commission’s preliminary commitment for immediate humanitarian assistance of 30 million euros in addition to 92 million euros in preliminary commitments by member states. It welcomed the commission’s preliminary commitment of 100 million euros for early, non-humanitarian assistance. For instance, in rehabilitation and reconstruction. And it took note of the commission’s preliminary indication that 200 million euros will be available for longer-term responses. It called for, in due course, and after post emergency needs to be fully assess, an International Conference on Haiti. In total, this is a big response in a short time. Assistance will be targeted too, to the Dominican Republic.

Finding enough money is going to be a less of a problem at the moment than getting it out to the people who need it. We must insure it’s properly targeted towards sustainable political and physical reconstruction. Following the United Nations secretary-general’s request for help to transport of humanitarian aid and for additional police units, the council invited me to identify EU member state contributions and to make proposals for this mobilization, and I’m working on this.

We convened a political and security committee meeting immediately after the council meeting to follow up. Early indications of member state contributions are already now available including possible contributions through the European Joint Air-Marine Force. Work will continue in the council proprietary groups during the coming days with a view to a rapid, targeted response.

The Foreign Affairs Council meet again next Monday. We’ll follow up on Haiti and consider further action. This is a major test of EU foreign policy in the new world of the Lisbon treaty. The people of Haiti, and our own people, expect a swift, effective, and coordinated response. This is what we believe we are delivering. I look forward to working with Parliament on this. I’m pleased to be here today to give you this information and to hear your views.

Source: https://webgate.ec.europa.eu/sr/speech/eus-aid-haiti
Our naked focus is on working with the United Nations of a Haitian leadership to ease the suffering of adhesion people a commitment to rebuilding Haiti is long-term many European citizens have also lost their lives and around 1000 I still on the country for who has responded quickly not seeking headlines but with the total focus on getting help to the people in need

Advice from the United Nations be resisted the urge to travel to Haiti immediately that would only have diverted attention away from the relief effort we will have Coast travel I-80 as soon as it is appropriate and I've agreed with commissioner to Gooch that he will travel to Haiti this week just send the eu's condolences on on July not commitment to the people he will also take the opportunity to review the most pressing me for the weeks or months to come

In the meantime we continue to work flashlight on all tracks humanitarian political and security over the past days I have been in constant contact with Secretary of State Clinton with the United Nations leadership withyou for a Minister's on with Canada which is leading the friends of Haiti group all this for the view to insure an effective and coordinated International response

I will travel to the United States this week to

เป้าหมายที่เปลือกายของเราคือการทำงานร่วมกับองค์การสหประชาชาติในการเป็นผู้นำของชาวฮิตีเพื่อบรรเทาความเดือดร้อนของคนยึดเหนี่ยวที่มุ่งมั่นในการฟื้นฟูฮิติเป็นเวลานานเขาจะไม่ใช้การเดินทางไปตามชายฝั่ง I-80 ทันทีที่เหมาะสมและยืนได้ตกลงกับข้าราชการของอังกฤษ ว่าจะเดินทางไปยังชายฝั่งในสัปดาห์นี้เพียงแค่ส่งความเสียใจของผู้ที่จะไม่ใช้โอกาสในการตรวจรอบมากที่สุดดังนั้นสำหรับสัปดาห์หรือเป็นเดือนที่จะมา

ในขณะเดียวกันเราดำเนินการฟื้นฟูให้ปลอดภัยทั้งหมดในด้านมนุษยธรรมทางการเมืองและความมั่นคงในช่วงหลายวันที่ผ่านมาแม้แต่คิดต่อไปกับเส้นทางการของรัฐบาลที่ต้องเน้นการฟื้นฟูหลังจากการเป็นผู้นำขององค์การสหประชาชาติกับรัฐมีสำคัญกว่าเราจะต้องเคลื่อนตัวไปสู่การคืนชีพที่จะมา

ฉันจะเดินทางไปสหรัฐอเมริกาในสัปดาห์นี้เพื่อ
follow up on this and other issues with the United States of ministration on with the UN secretary-general and others in New York do you learn has asked for Urgent financial aid 575 million u.s. dollars to transport humanitarian Aid yesterday the United Nation secretary-general ask her please and Military reinforcements of the UN peacekeeping Mission as well.

Don't mobilize on coordinate our response I have to Spanish Tu presidency to call an extraordinary Foreign Affairs canceled yesterday for the first time since the Lisbon treaty engine true fact we're not putting together the efforts of the commission on the console Secretariat on the member states in a comprehensive approach onto my overall coordination this is groundbreaking we had a productive on the need for rapid response with the United Nations.

In terms of financial assistance the council reach the following conclusion it's welcome to the commission's plenary commitment for mediate humanitarian assistance of 30 million euros in addition to 93 million euros in preliminary commitments by numbers face welcome to the commitment of 100 million euros in Rehabilitation and reconstruction.

It's welcome to the commission's preliminary commitment of 100 million euros early non humanitarian Assistance or entrance instance in Rehabilitation and reconstruction and it
took note of the commission Stella Mae indication that's $200 in euros will be available for longer-term responses

It called for intercourse after post emergency needs to be fully assess and International Conference on Haiti in total this is a big response in a short time assistance will be targeted to the Dominican Republic

Finding enough money it's going to be that's our problem at the moment and getting it out to the people who need it we must Insurance property targeted sustainable political and physical reconstruction United Nation secretary-general request for help to transport of humanitarian agent for additional police units and invited me to identify contributions and to make proposals for this mobilization and I'm working on this

be convened a political and security committee meeting immediately after the council meeting to follow up early indications of Memphis take contributions are already now available including possible contributions for the European John Danbury Force what could continue in the coming days with a response

The Foreign Affairs Council meet again next Monday will follow up on Haiti and consider for the action this is a major test of the uniform policy in the new world of the Lisbon treaty the people of Haiti and our own people
expect a swift effective and coordinated response this is what we believe we are delivering I look forward to working with Parliament on this I'm pleased to be here today to give you this information on to hear your views
Appendix C

Sample B1 Transcript

“European Road Safety Day 2007”

I have now pleasure, we have a little bit of work to do with you, but I think it was still be very varied and interesting. We have people from all over Europe and the world tell different aspects of what could interest young people in traffic, and so we start without further ado, but I still have one announcement to make on the Working Session One on alcohol and drugs in traffic.

Before we start, again a reminder the Red Cross is saying they have still capacities for you to refresh your first aid skills. You just go there, you don't need to be inscribed, if are inscribed it's good, if not you just go there and they'll be happy to teach you one or two tricks. Take advantage of it, it's really something which I think is this useful and of course free of charge.

Okay, I give the floor to mister Apfel from the UK.

Thank you very much, this session is going to focus on the risk factors related to alcohol and drugs. We know that drivers 16 to 20 have three times three times more likely to crash than drivers older than 30 when at any alcohol level. And use of recreational drugs, particularly in young men, also greatly increases their crash ability. As we heard this morning and saw in this very useful report card on member states that there's great differences that exist between countries in Europe. That leads us to believe that there's great opportunities to make changes and improve and reach the goals that have been set. What kind of actions have been taken in the alcohol and drug behavioural modification risk area? Getting training is quite important, and I got to say that the greatest gift that the UK government has given me was to fail my son twice in his driving test. It took him over a year to pass his driving test. He was a much better driver after a year.

Strong and rigid restrictive blood alcohol level settings. The European Parliament and WHO recommend zero-level for novice drivers. But not just setting the levels, enforcing the laws is very important. High visible random breath testing associated with the reductions in drunk, drink-driving. Also control on sales and pricing, important. Controls on advertising, certainly we saw in the statement by the youth that they were concerned about this, where they call upon all private companies to refrain from showing advertisements or TV shows that glamorize speeding and drink-driving during times with youth might be watching. I might say that's 24/7 actually.
We found that statutory controls on advertising work better than voluntary controls. And most important really raising awareness. And this meeting, this week, these days I think have been a tremendous platform to do just that. And on that platform standing with peer-to-peer counselling and bringing the young people to talk with young people, most important.

That’s what we going to present to you this afternoon. We’re going to talk a little bit about why of alcohol and drugs is important to focus on, and then get into looking at some, how it is being done in different parts of Europe. I would like to start the afternoon session by inviting Eileen Verstet to present information on the effects of alcohol and behaviour. Dr. Verstet is from the university of Ghent; the floor is yours.

Online now pleasure with bit of work to do with you but I think it was still be very very that interesting we have people from all over Europe in the world tell different aspects of what could interest young people in traffic and so we start without further Ado but I still want to know it's been to make on The Working session one I'll go and relax in traffic before we start again.

Begin a reminder the Red Cross is saying they have still capacities for you to refresh your first aid skills you just go there you don't need to be inscribed if you unsubscribe is good if not you just go there and I'll be happy to teach you one or two drinks take advantage of it it's really something which I think is this useful and of course free of charge.

Okay I guess. Phil from the UK thank you very much this session is going to focus on the risk factors related to alcohol and drugs we we know that drivers 16 to 20 have three times three times more likely to crash than drivers older than 30 win at any alcohol level and use of recreational drugs that particular and young men also greatly increases their crash ability as we heard this morning and Soarin report card on member states that there's great differences that exist between countries in Europe that leads us to believe that there's great opportunities to
make changes and improve and reach the goals that have been said things have been taken in the alcohol and drug behavioral modification risk area getting training is quite important and I got to say that the greatest gift that the UK government has given me was to fail my son twice in his driving test it took him over a year to pass his driving test it was a much better driver after a year

Strong and rigid restrictive blood alcohol level of the settings the European Parliament to whom recommend zero level for novice drivers but not just setting the levels enforcing the laws very important High visible random breath testing associated with the reductions in drunk drunk driving also control on sales and pricing important controls on advertising on a certainly we we soar in the statement by the youth that they were concerned about this where they call upon will private companies to refrain from showing advertisements or TV shows that glamorize speeding and drink driving during times with youth might be watching I might say that’s 24/7 actually

I found that statutory controls and advertising work better than voluntary controls and most important really raising awareness and this meeting this week these days I think I’ve been a tremendous platform to do just that and on that platform standing with peer-to-peer counseling and bringing the young people to talk with young people most important
That's what we going to present to you this afternoon we're going to talk a little bit about why of alcohol and drugs is important to focus on and then get into looking at some how it is being done in different parts of Europe would like to start the afternoon session by inviting Ally invest it to present this information on the effects of alcohol and behavior towards yours.
Appendix D

Sample B2 Transcript

“Beyond GDP: Measuring Progress, True Wealth, and the Well-being of Nations”

I think that I draw three main conclusions from this conference. The first one is about the actors. We have seen these days a lot of people representing different constituencies with a very similar view not just to go beyond GDP but to develop measures that can be used, not by just policymakers, but also by other decision-makers including citizens. Including ourselves as individuals.

We heard about civil societies’ initiatives like local communities associations who have developed measures of progress, of well-being, and so. We heard the about the commitment of some governments to take this forward. We heard the representatives of parliament, not just European parliament, to associate themselves to this effort. We heard the international organizations, we heard statisticians, and also we had some representatives of the business world, media, and citizens.

How to bring these different constituencies all together? This is a big Challenge and is one of the points that I would like to suggest for our discussion this afternoon.

But it’s clear that we have a lot of actors involved and committed to take this agenda forward, and the risk is that we increase confusion instead of increase cooperation and collaboration. We have to find a way to work together and give a sense to the final stakeholders that we are working together and not fighting each other; as somebody said today that our word that is too much oriented to compete and in some cases to fight.

Second point: measures. We have heard about the different approaches and my personal account, conclusion, and summary, again for discussion this afternoon, is the following: Single number? No, it’s not a solution. We cannot reduce the whole complexity of our world to a single number. GDP is not the right single number, but there is any other single number that can really represent the complexity of our world?

We heard the sum of positive words about Composite Indicators and I think that we should share this view. Composite indicators can be helpful, do not answer all questions, but it can help. In the short run, Key Indicators were quoted as a possible solution with the possibility of giving good
answers, with some weaknesses, but I heard from many speakers that key indicators are used and they are seen as a possible way forward. The long-term solution, and we have to accelerate as much as possible this transition, is to extend the Economic National Accounts Framework to other domains. We have done a lot on the environmental side; the issue here is mainly implement what has been developed. On the social side human capital or, according to others, Social Capital we are still lagging behind. But it’s clear that in the medium-long-term, the extension of national accounts seems very promising although costly.

Final point: it’s not enough to engage people, it’s not enough to develop good measures, we need to invest on communication and empowerment. These two key words were mentioned several times. We need to engage people. We need to help people to understand the world which they’re living and their society. We need to increase accountability of governments. We need to invest, like in tools like those that Hans Rosling show to us, to allow people to transform information into knowledge. Again this is a big effort, it’s not just about statistics, it’s an investment in public good, which is common knowledge.

I stop here my intervention. I will come back a little bit later on some proposals for the way forward, but I want to give now the floor to the first speaker this afternoon, and Walter is very well known, chief statistician, but also he has been working a lot for many years on green accounts and so his heart, let’s say, is very close to what we have been debating these days.

I think that I drove three main conclusions from this conference the first one is about the actors who have seen these days a lot of people representing different constituencies with a very similar you know just to go beyond that can be used by policymakers but also by other decision-makers including citizens including ourselves as individuals

we heard about secret societies initiatives like local communities associations who have developed measures of a progress of well-being and song we heard the about the commitment of some governments to take this forward we heard the representatives of parliament or just you parliament to associate themselves to decipher. We heard the international organizations we heard statisticians and also we had some representatives of the business world media and citizens

How to bring this different constituencies all together this is a big Challenge and is one of the points that I would lie to suggest for our discussion this afternoon

But it’s clear that we have adult actors involved and committed to take this agenda Ford and that we increase confusion inside of a increase Corporation collaboration we have to find a way to work together and The Giver sense to the final stakeholders that we are

 ผมคิดว่าผมได้ผลักดันสามข้อสรุปหลักจากการประชุมครั้งนี้เรื่องแรกคือเรื่องเกี่ยวกับนักแสดงที่ได้เห็นในวันนี้มีผู้คนจำนวนมากที่เป็นตัวแทนของเขตเลือกตั้งที่แตกต่างกันซึ่งคล้ายคลึงกันมากที่คุณรู้เพียงเพื่อไปไกลกว่าที่ผู้กำหนดนโยบายสามารถใช้ได้ผู้มีอำนาจตัดสินใจรวมถึงประชาชนรวมถึงตัวเราเองด้วยเช่นกัน

เราได้ยินเกี่ยวกับการริเริ่มของสังคมลับเช่นสมาพูธชุมชนในโครงการที่พัฒนามาตรการความก้าวหน้าและสวัสดิการเพลิดเพลินได้อยีเกี่ยวกับความมุ่งมั่นของรัฐบาลบางแห่งในการดำเนินการล่าช้าต่อไปนี้ได้ยินจากตัวแทนของรัฐบาลหรือเฉพาะรัฐบาลที่จะเชื่อมโยงต่อเอง แพร่หลาย เราได้ยินจากองค์กรระหว่างประเทศที่เราเคยได้ยินเสียงและเรามีส่วนทางจากเสียงและเพลงของสิ่งต่างๆ

วิธีการที่จะนำเขตเลือกตั้งที่แตกต่างกันนี้เข้าหัวขึ้นกันนี้คือความท้าทายที่ยิ่งใหญ่และเป็นจุดที่ฉันจะไปหาเพื่อแนะนำให้เราธุรกิจในบ่ายวันนี้

แต่เห็นได้ชัดว่าเรามีนักแสดงผู้ใหญ่ที่เกี่ยวข้องและมุ่งมั่นที่จะใช้ชีวิตนี้กับพอร์และเราเพิ่มความสัมผัสในการทำงานร่วมกันของบริษัท ที่เพิ่มขึ้นเราร้อนทางวิธีที่จะทำงานร่วมกันและความรู้สึกของผู้ที่ยังไม่มีส่วนได้เสียผู้ดูที่เรากำลังทำงานร่วมกันและไม่ต้องถูกเป็นบางคนสำหรับวันนี้ว่าคำของเรามาก
working together and not fighting each other as somebody said today that our word that is too much oriented to compete and in some cases to fight.

Second point measures we have heard about the different approaches and my personal account conclusion summary again for discussion this afternoon is the following single number know we can reduce the complexity of a single number single number but there is any other single number that can really represent the complexity of our world.

We heard the sum of positive words about composite indicators and they think that we should share this view Conquest indicators can be helpful do not answer all questions that I can help in the short run teenagers were quoted as a possible solution with the possibility of giving good answers with some weaknesses that I heard from many speakers that key indicators are used and are seen as a possible way forward the long-term solution and we have to accelerate as much as possible this transition is to extend the economic national accounts framework to other domains we have done it all on the environment outside the issue here is mainly Implement what has been developed on the social side human capital or Social Capital we are still lagging behind and but it’s clear that in the middle of the extension of national accounts seems very promising although costly.
Final point is not enough to engage people is not enough enough to develop good measures we need to invest on communication and empowerment this two key words were mentioned several times we need to engage people we need to help people to understand the words which their living and their society we need to increase accountability of government we need to invest like in to slide those that Hans rosling show to us to allow people to transform information into knowledge again this is a big effort is not just about statistics is an investment in public good which is common knowledge

Ice computer my intervention I will come back a little bit later on some proposals for the way forward but I want to give now the floor to the first Speaker this afternoon and Valtor is very well known priest at the station but also he has been working a lot for many years and green accounts and so his heart is very close to what we have been debating this days