Comprehension and Trust in Crises: Investigating the Impact of Machine Translation and Post-Editing

Alessandra Rossetti[(1,2)], **Sharon O'Brien**[(1,2)], **Patrick Cadwell**[(2)]
[(1)] ADAPT Centre
[(2)] School of Applied Language and Intercultural Studies
Dublin City University, Dublin
Ireland
{alessandra.rossetti, sharon.obrien, patrick.cadwell}
@dcu.ie

Abstract

We conducted a survey to understand the impact of machine translation and post-editing awareness on comprehension of and trust in messages disseminated to prepare the public for a weather-related crisis, i.e. flooding. The translation direction was English–Italian. Sixty-one participants—all native Italian speakers with different English proficiency levels—answered our survey. Each participant read and evaluated between three and six crisis messages using ratings and open-ended questions on comprehensibility and trust. The messages were in English and Italian. All the Italian messages had been machine translated and post-edited. Nevertheless, participants were told that only half had been post-edited, so that we could test the impact of post-editing awareness. We could not draw firm conclusions when comparing the scores for trust and comprehensibility assigned to the three types of messages—English, post-edits, and purported raw outputs. However, when scores were triangulated with open-ended answers, stronger patterns were observed, such as the impact of fluency of the translations on their comprehensibility and trustworthiness. We found correlations between comprehensibility and trustworthiness, and identified other factors influencing these aspects, such as the clarity and soundness of the messages. We conclude by outlining implications for crisis preparedness, limitations, and areas for future research.

1 Introduction

Societies are becoming increasingly multicultural and multilingual, mainly as a result of economic migration and displacement (O'Brien and Federici, 2019). In Ireland, for example, there are more than 500 thousand non-Irish nationals, the majority of whom come from a country where English is not the official language, e.g. Poland, Lithuania, Brazil, and Italy (Central Statistics Office, 2016). Non-native speakers of a language—and especially those with limited proficiency—need to overcome considerable communication challenges in the contexts of crises (Santos-Hernández and Morrow, 2013; Sherly et al., 2015).

Taking again Ireland as an example, flooding is the most common hazard that the country needs to manage (Jeffers, 2011). When substantial, flooding poses a threat to infrastructure, business, and also people's health (Major Emergency Management, 2016). In order to be safe and act upon the messages sent by emergency responders, linguistically diverse communities need to be able to comprehend and trust those messages (Alexander and Pescaroli, 2019). Machine translation (MT) and post-editing (PE) can play a role in crisis communication but their application needs careful consideration.

This paper describes the results of a survey whose goal was to address two important gaps in relation to the role of MT and PE as enablers of multilingual communication in crises. Specifically, we set out to gather empirical evidence on the impact of MT and of PE awareness on comprehension of and trust in messages disseminated by emergency responders to prepare the public for a

specific weather-related crisis: flooding. The translation direction under analysis was English to Italian (see Section 3 for our research questions). The choice of this translation direction was motivated by the substantial number of native speakers of Italian living in English-speaking countries where flooding is common, such as the United Kingdom and Ireland (Central Statistics Office, 2016).

It is worth underlining the lack of clear distinctions between the concepts of *crisis*, *emergency*, *disaster*, or *hazard*. For the purpose of this study, we adopted a broad definition of *crisis*, understood as a non-routine and disruptive event, that poses a threat, and that usually involves the phases of preparation, response, and recovery (Alexander, 2002; Cadwell et al., 2019).

The remainder of this paper is organized as follows: Section 2 reviews and summarizes related work on MT, PE, comprehension, and trust, with a special focus on crisis contexts. Section 3 presents our research questions and the methodology that we adopted in order to answer them. Section 4 reports on the results of our survey, which are then discussed in Section 5, along with implications, limitations, and avenues for future research.

2 Related Work

Translation of crisis information into the first language of the target audience facilitates comprehension, as has been shown, for example, in the case of the 2014 Ebola outbreak (O'Brien and Cadwell, 2017). However, the importance of translation in crises is still either not acknowledged or discussed only superficially in policy documents and institutional checklists (O'Brien et al., 2018; O'Brien and Federici, 2019). This is surprising when considering that misunderstandings due to lack of translation have often resulted in increased vulnerability and loss of lives (Santos-Hernández and Morrow, 2013; Alexander and Pescaroli, 2019).

In addition to comprehension, the language in which information is conveyed can influence trust in the message, particularly in crisis situations (Translators without Borders, 2019). Previous research on trust, translation, and crises has mainly focused on how translation influences reasoning about trust among people affected by a crisis (Cadwell, 2015), with trust emerging as one of the challenges in the communication efforts of humanitarian organisations, along with low literacy levels and cultural sensitivity (Federici et al., 2019).

In crisis situations, MT has been a component of some communications, as shown, for instance, during the Haiti earthquake (Lewis, 2010) and, more recently, in refugee settings (Translators without Borders, 2016). MT is particularly helpful when large quantities of texts need quick translations into multiple languages (Cadwell et al., 2019). The utility of MT in crisis settings involving low-resource languages has also been empirically tested (Cadwell et al., 2019).

The relationship between MT and trust has received some attention since machine-translated outputs are far from flawless and fully accurate, even after the quality improvements introduced by the neural paradigm (Toral et al., 2018), thus often requiring PE. Research has revolved around approaches to identify machine-translated words, sentences or documents that pass a predetermined quality threshold and are therefore more trustworthy (Soricut and Echihabi, 2010).

The availability of these confidence, or trust, scores seems to be welcomed by translators (Moorkens and O'Brien, 2013), but the scores should be accompanied by an explanation of how they were obtained (Cadwell et al., 2017). Attention has also been given to the level of trust that professional translators attribute to machine-translated outputs and specific MT engines (Guerberof, 2013; Teixeira, 2014; Cadwell et al., 2017). Furthermore, lack of trust in MT has emerged as one of the reasons for its non-adoption among language service providers (Porro Rodríguez et al., 2017). Previous works have also focused on students, with mixed results—from a general lack of trust (Koponen, 2015; Briggs, 2018), to a tendency to almost uncritically trust the output (Depraetere, 2010).

More relevant to our research, a limited number of studies have focused on end users of MT—who often read translations for gist understanding (Specia and Shah, 2018)—and on their reliance on MT to locate information on websites (Gaspari, 2007), as well as on their tendency to use MT to translate from languages or documents of which they already have some knowledge, which might indicate a lack of complete trust in the output (Nurminen and Papula, 2018).

Research has also focused on the broader areas of acceptability, usability, readability, and com-

prehensibility of machine-translated texts among end users, and on how these aspects are influenced by different PE levels (Castilho and O'Brien, 2016; Screen, 2019). However, most of the research so far has focused on technical documents.

Accordingly, there is a lack of empirical evidence on: (i) the potential benefits of MT (as opposed to lack of translation) for end users' comprehension of and trust in crisis communication; and (ii) the potential impact on comprehension and trust of being aware that crisis messages have been post-edited. We set out to fill these research gaps.

3 Methodology

3.1 Research Questions

Having in mind the research gaps outlined in Section 2.2, we conducted a survey to address the following research questions (RQ):

RQ1. What is the impact of machine translation on comprehension of and trust in messages disseminated to prepare the public for a weather-related crisis?

RQ2. What is the impact of post-editing awareness on comprehension of and trust in messages disseminated to prepare the public for a weather-related crisis?

As specified in Section 1, the translation direction under analysis was English to Italian.

3.2 Survey Setup and Circulation

All of the survey questions and instructions were in Italian. The survey received approval from Dublin City University Research Ethics Committee (DCUREC/2019/209). It was preceded by a plain language statement and an informed consent form (also in Italian) describing the research in lay terms for the participants.

Initially, the survey targeted native speakers of Italian living in English-speaking countries, as they would represent a realistic audience for crisis messages delivered by emergency responders in English. However, an initial analysis of the responses from this pool of Italian participants showed that their self-reported level of English was very high (Section 4.1). Accordingly, to gather data from Italian speakers with lower levels of English proficiency—thus gaining a broader range of perspectives—we also circulated a slightly modified version of the survey among native speakers of Italian living in Italy (see Section 3.3 for details on the slightly modified version). These participants were also a realistic audience considering the high number of Italians who travel from Italy to English-speaking countries for tourism, school- or business-related purposes (Tourism Ireland, 2018).

The survey in both its versions was circulated online through word-of-mouth; social media; and newsletters from universities, Italian embassies, and organisations promoting Italian culture in English-speaking countries (from the United States, to Ireland, to New Zealand).

3.3 Survey Structure and Experimental Design

The survey began with two questions to check participants' eligibility, namely: (i) that their native language was Italian; and (ii) that they lived in an English-speaking country. In the version of the survey targeting Italians in Italy, the second eligibility question was not present.

The survey then continued with a series of questions on the participants' demographic characteristics and background, namely their age, gender, self-reported level of English proficiency, frequency of use of English, familiarity with MT systems, and reasons for their use. With regard to the questions on self-reported English proficiency and on the frequency of use of the English language, these questions were taken from Anderson et al. (2018), and they involved asking participants: (i) to rate their English conversation, writing, reading, and listening skills on a scale from 1 (low) to 5 (high); and (ii) to indicate how often they spoke, wrote, listened, and read in English. Native speakers of Italian in English-speaking countries were also asked about how much time they had lived abroad, and the frequency of flooding in their country of residence (Section 4.1).

The participants were subsequently presented with information and instructions regarding the experimental tasks. Specifically, they would first be shown three messages dealing with preparation for a flooding crisis: one message would be in English, while the other two would be Italian translations of two different messages. They were also told that, of the two translations, one had been produced by Google Translate and had not been corrected by anyone, while the other had also been produced by Google Translate but then corrected by a native speaker of Italian. We

used *corrected* (rather than *post-edited*) because our participants might not have been familiar with the concept of PE. We also specified that we would let them know which MT output had been post-edited/corrected beforehand.

At this stage, we used deception since both machine-translated messages had actually been post-edited by the first author (see Section 3.4 for details on PE level). We used deception for two reasons. First, if we had not post-edited one of the two machine-translated messages, we would have introduced MT quality as a confounding variable—in other words, the different quality of the two machine-translated messages would have been likely to influence comprehensibility and trust scores. By post-editing both outputs, we ensured quality was comparable, and this allowed us to determine whether awareness of PE in itself influenced scores of comprehensibility and trust given by end users. Secondly, due to the critical nature of the messages, we deemed it risky to circulate unedited content with potential errors.

We adopted a within-subjects design whereby, for each of the three messages (one in English and two Italian translations), each participant was instructed to answer the following questions:

(i) How much do you trust this message on a scale from 1 (don't trust it at all) to 4 (trust it completely)?

(ii) How likely are you to comply with these instructions on a scale from 1 (very unlikely) to 4 (very likely)?

(iii) How comprehensible do you find this message on a scale from 1 (totally incomprehensible) to 4 (easily comprehensible)?

All participants read and evaluated the same messages, and each message was always seen in the same condition. We added a question on compliance as an additional measure of trust (Liu et al., 2018). We used four-point scales to avoid mid-point bias. For each of the three questions, participants were also given the option to explain the reasons behind their scores as answers to open-ended questions. Finally, after reading and scoring the first set of three messages, participants could either conclude the survey, or read and evaluate a set of three more messages. To counterbalance a potential fatigue effect, the order in which the English message and the two Italian translations were presented to participants varied between the first and the second set of messages, but not within set.

3.4 Experimental Materials

We took the crisis preparedness messages from the Irish website *Be Winter Ready*.[1] The PE applied to the machine-translated messages can be classified as full PE since we aimed to produce outputs that were both fluent and accurate (TAUS, 2010). Average BLEU score based on comparisons between raw and post-edited messages was 55.76. However, as the extracts in Section 4.2 show, a few participants believed that the fluency could have been improved further.

Since the readability level of the English source messages—both the one that we kept in English and the ones that we machine translated into Italian—might have represented a confounding variable influencing comprehensibility scores, we selected messages with a similar or almost similar readability level. Specifically, according to the Flesch-Kincaid Grade Level formula, all English messages could be understood by readers between 11 and 16 years of age.

To further ensure comparability, the three messages in each of the two sets (Section 3.3) began with the same introductory sentence. The three messages in the first set all began with "If you find that you are in a flood prone area, there are a number of steps that you can take to make your property more resilient to flooding. For example…", as they dealt with property protection. On the other hand, the three messages in the second set revolved around people protection and began with the introductory sentence "If you find that you are in a flood prone area, there are a number of steps that you can take. For example…". These introductory sentences were then followed by specific instructions, such as "Assess if your property is at risk from flooding" in the first set, or "Have medication to hand (if needed)" in the second set. To avoid a learning effect, the three instructions in each set were different.

4 Results

4.1 Participants' Background

A total of 61 participants took part in the survey. All the participants were native speakers of Italian, with 48 of them living in an English-

[1] The *Be Winter Ready* website is available here: https://www.winterready.ie/en

speaking country and 13 living in Italy. Most participants were aged between 29-39 (46%), followed by participants aged 40-50 (29%). We achieved good balance between male (52%) and female (46%) participants—2% of the participants did not specify their gender.

Among the 48 participants based outside Italy, most of them reported having lived in an English-speaking country either between five and ten years (N=13), or between ten and 20 years (N=13), with seven also stating that they had lived in an English-speaking country for more than 20 years. Unsurprisingly, when asked to self-report their level of English proficiency in terms of conversation, reading, writing, and listening, most participants within this cohort reported five out of five. Furthermore, the vast majority of them stated that they spoke, wrote, read, and listened in English either always or most of the time.

In contrast, most participants based in Italy reported having a lower level of English proficiency—most of them selected one (out of five) to rate their English conversation skills, and three (out of five) to rate their listening, writing, and reading skills. In line with these scores, most of the participants based in Italy stated that they spoke, listened, and wrote in English only rarely. However, most of them reported reading in English sometimes. In other words, our two cohorts of participants—namely, Italians living in English-speaking countries and Italians living in Italy—were different enough in terms of English proficiency, which allowed us to gather data from a broad range of potential users of crisis communications (Section 4.2).

42% of the 48 participants living in an English-speaking country stated that flooding—namely, the weather-related crisis that is the focus of our study—was common where they lived, with 14% not knowing, as shown in Figure 1.

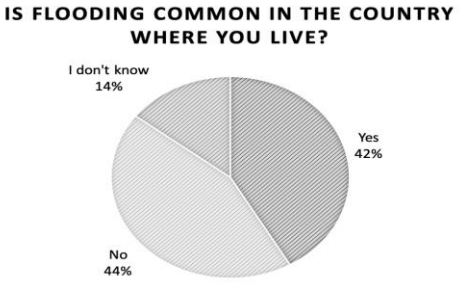

Figure 1. Percentage of participants (not) familiar with flooding

With regard to the use of MT systems, of all the 61 participants, 48 reported using MT systems. The reasons for their use of MT are reported in Figure 2, where the number of selections is higher than the number of participants because participants could select more than one option. Assimilation was the most common reason, followed by dissemination. This result was relevant as it showed that these end users could potentially use MT to translate crisis messages delivered in a language with which they were not familiar.

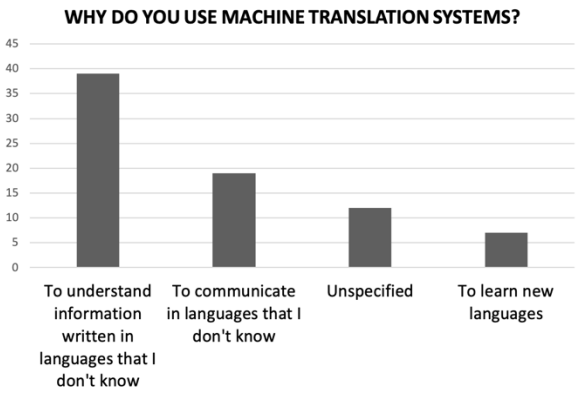

Figure 2. Participants' reasons for use of MT

4.2 Comprehensibility and Trust

The tables below contain descriptive statistics—mean and standard deviation (SD). Table 1 reports the comprehensibility scores. Table 2 contains the trust scores, and Table 3 shows the trust as compliance scores. In each table, we first reported the scores provided by all 61 survey participants combined, and then by Italians living in English-speaking countries and by Italians living in Italy separately, as these two groups differed substantially in terms of English proficiency (Section 4.1). We combined scores assigned by participants to both sets of messages (Section 3.4). In the interests of clarity, in the tables and elsewhere in this paper we used *raw messages* for those MT outputs that had also been post-edited even though participants thought that they had not been—our deception condition (Section 3.3). The highest scores are highlighted in bold.

With regard to comprehensibility (Table 1), it can be observed that: (i) the messages labelled as post-edited received the highest average scores by all three cohorts of participants; (ii) participants living in Italy—and having a lower level of English proficiency—seemed to benefit more from the translations labelled as raw, compared with the English messages, than participants living in English-speaking countries. As far as trust

is concerned (Table 2), results were more varied: (i) the messages labelled as post-edited were not associated with highest average scores; but again (ii) differently from participants in English-speaking countries, participants living in Italy showed higher trust in the messages labelled as raw, compared with the English messages. With regard to trust measured in terms of compliance (Table 3), we observed that, regardless of their level of English proficiency, participants showed higher compliance with the message in English, compared with the Italian translations. It should be noted, however, that the differences in scores reported in Tables 1-3 are slight, and a series of repeated measures ANOVAs run in SPSS found these differences to be not significant ($p>.05$).

	Comprehensibility		
	English messages	**Raw messages**	**Post-edited messages**
Total participants (N=61)	3.45 (.83)	3.54 (.75)	**3.64 (.64)**
Italians abroad (N=48)	3.66 (.62)	3.64 (.63)	**3.74 (.51)**
Italians in Italy (N=13)	2.71 (1.04)	3.18 (1.01)	**3.29 (.92)**

Table 1. Comprehensibility scores

	Trust		
	English messages	**Raw messages**	**Post-edited messages**
Total participants (N=61)	**3.36 (.80)**	3.29 (.82)	3.35 (.90)
Italians abroad (N=48)	**3.49 (.74)**	3.34 (.77)	3.46 (.78)
Italians in Italy (N=13)	2.88 (.85)	**3.12 (.99)**	2.94 (1.19)

Table 2. Trust scores

	Trust (compliance)		
	English messages	**Raw Messages**	**Post-edited messages**
Total participants (N=61)	**3.53 (.75)**	3.35 (.90)	3.38 (.95)
Italians abroad (N=48)	**3.67 (.59)**	3.46 (.80)	3.56 (.78)
Italians in Italy (N=13)	**3.00 (1.0)**	2.94 (1.14)	2.76 (1.25)

Table 3. Compliance (trust) scores

Using SPSS software, we also examined potential correlations between comprehensibility scores and trust scores. The results, reported in Table 4, showed that comprehensibility scores and trust scores had a statistically significant linear relationship for all three types of messages ($p<.01$). The direction of the relationship was positive, and the strength of this association went from moderate to fairly strong ($.5 < r_s < .7$). In other words, regardless of how the messages were labelled (i.e. raw MT vs. PE) and regardless of translation, greater comprehensibility was often associated with greater trust.

	Trust	**Trust (compliance)**
Comprehensibility	*English messages* .69*	*English messages* .66*
	Raw messages .53*	*Raw messages* .66*
	Post-edited messages .55*	*Post-edited messages* .62*

Table 4. Results of the Spearman Correlation[2]

The qualitative data collected through the open-ended questions in the survey (Section 3.3), and coded with the NVivo software, complemented these scores and guided their interpretation. We used thematic analysis (Braun and Clarke, 2012) to identify the main reasons behind the comprehensibility and trust scores that the participants assigned. Our analysis identified seven themes in the participants' responses, namely: clarity; soundness; helpfulness; fluency; style; source; and individual differences.

Figure 3 shows how many times each reason was mentioned per message and per each object of investigation among native Italian speakers living in English-speaking countries. Figure 4 reports the same data for the cohort living in Italy. Again, we counted and analysed the answers given by the participants when evaluating both sets of crisis messages (Section 3.3). Participants could indicate more than one reason for each of their scores.

In line with the moderate to fairly strong correlations in Table 4, Figures 3 and 4 show that clarity (defined as simplicity and comprehensibility of language) was regarded by numerous participants as a reason to trust the messages. For participants living in Italy and having lower English proficiency, clarity was needed to trust the messages particularly when the messages were in English, which might explain the slightly lower average score that they assigned to the trustworthiness of English messages (Table 2).

[2] Statistical significance (*) is at the .01 level.

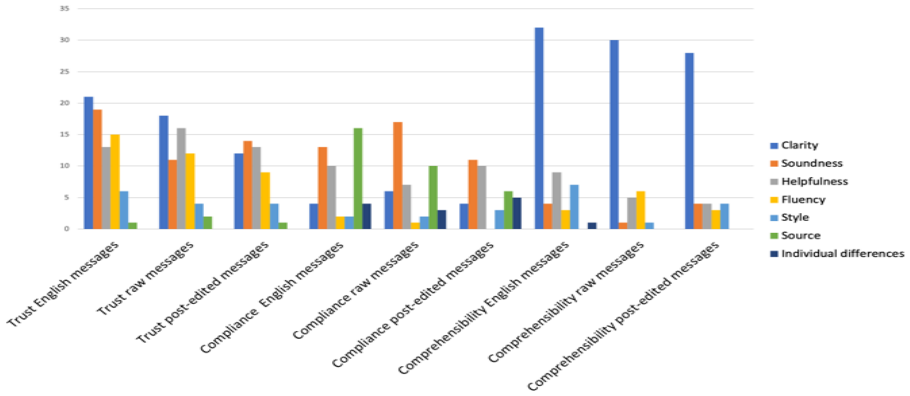

Figure 3. Mentions of themes by participants in English-speaking countries, cross-referenced with experimental conditions

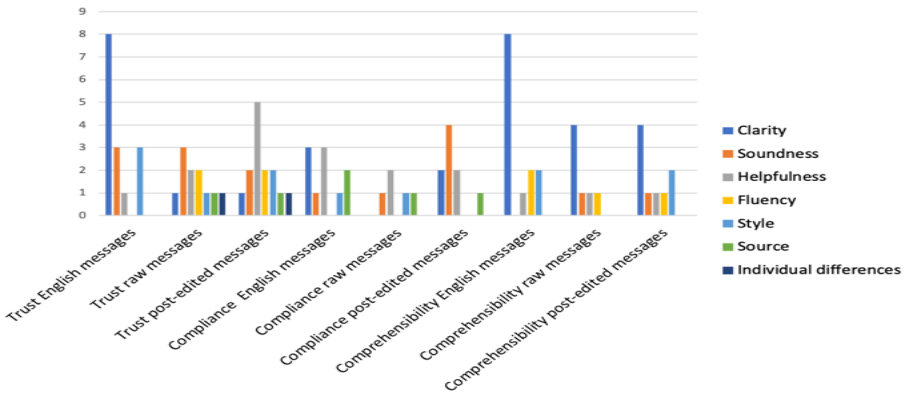

Figure 4. Mentions of themes by participants in Italy, cross-referenced with experimental conditions

As would be expected, clarity also emerged as a common reason influencing comprehensibility scores. A few participants mentioned the features that rendered a message clear, such as the absence of technical terms, simple noun and verb phrases, or the use of common words. It should be remembered that our experimental materials could be understood by readers between 11 and 16 years of age (Section 3.4).

When evaluating their level of compliance, clarity seemed to be less relevant to participants. In contrast, the soundness, the helpfulness, and the source of the messages seemed to be determining factors. Often, the soundness and helpfulness of the messages also determined the participants' level of trust in the messages. See, for examples, extracts below[3] in which participants explain why they would trust and comply with a specific crisis message:

P10: Logical and reasonable explanation.

P20: It's reasonable, and the task can be carried out easily, and it's for my benefit.

P09: In general, if it [the message] comes from emergency managers, it means that the information provided is accurate.

Another aspect occasionally influencing trust seemed to be style—this theme included the tone and register of the message. Specifically, several participants mentioned the authoritative tone, the directness of the message, and the sense of competence emerging from the messages—especially those in English—as reasons to trust them.

Individual differences, and especially previous experience of a weather-related crisis, also emerged as a reason for compliance among Italians living in English-speaking countries. This is not surprising considering that almost half of them reported living in a country where flooding is common (Figure 1).

Considering the specific focus of this paper on the impact of MT and PE (Section 3.1), as well as on how messages were labelled, it is interesting to notice that the theme of fluency—capturing participants' mentions of how (un)natural the language of the translated messages was—emerged as one of the reasons behind participants' trust and comprehensibility

[3] The answers in Italian were translated into English by the first author.

scores. See, for instance, the following explanations for assigning a specific score:

P14: The translation is correct, but it could be improved.

P22: Even though the message is clear, it's obvious that it's a raw translation.

P26: Message clear and simple, with no errors.

P05: Because it contains no errors, and you can't tell that it's an automatic translation.

P07: The message has been translated clearly and correctly, with no obvious grammar or syntactical errors.

P27: Convoluted, not fluid.

Despite these mentions of translation, Figure 3 and Figure 4 show that the other themes—and especially clarity, soundness, helpfulness, and source—had a stronger impact on participants' reported comprehensibility and trust. Interestingly, this observation on the somewhat lower impact of translation, and of how translation was labelled, is in line with the slight differences reported in Tables 1-3 between the scores assigned to English messages, to messages labelled as raw MT, and to those labelled as post-edited. Furthermore, participants' responses to the open-ended questions seemed to cluster around the same themes depending on whether the question was on trust, compliance, or comprehensibility, and regardless of whether they were reading the English message, the purported raw translation, or the post-edited translation. For instance, the importance of the source (i.e. emergency managers) was mentioned by several participants when indicating the reasons behind compliance, but was absent when they discussed their comprehensibility scores (Figures 3-4).

5 Discussion and Conclusions

With this survey, we set out to investigate the impact of MT and PE awareness, in the English to Italian direction, on comprehension of and trust in messages disseminated to the public in the context of preparation for a specific weather-related crisis, i.e. flooding (Section 3.1).

Overall, we found slight and non-significant differences in terms of scores between English, purported raw, and post-edited messages. However, some interesting trends emerged, namely: (i) some beneficial effect of MT on comprehension and trust among end users with low English proficiency; (ii) a tendency to comply more with messages in English, possibly as a result of their authoritative tone/style (Section 4.2); and (iii) labelling of messages as *post-edited* resulting in some improvement in comprehension, but not in trust. The absence of a beneficial effect of PE awareness on trust might be due to: (i) the purported MT outputs having also been post-edited and, therefore, appearing equally trustworthy; (ii) the fact that the fluency of the declared post-edits could have been improved further (Section 4.2). In line with these points, the fluency of the translations had some impact on how comprehensible and trustworthy the messages appeared to be to our participants. Interestingly, after comparing the influence of different PE levels among end users, Van Egdom and Pluymaekers (2019) found that full PE led to positive judgements in terms of language use and style, but did not result in a significant improvement of the perceptions (including trustworthiness) of the sender of a product.

Regardless of how the messages were labelled, several aspects of crisis messages were recognized as particularly important by participants, especially the clarity, the soundness, the helpfulness, and the source of the messages. The importance of clarity for comprehension could be expected. On the other hand, results regarding trust are particularly interesting as they align with models of trust (Mayer et al., 1995) according to which the decision to trust is determined by, among others: the competence of the trustee (e.g. their ability to provide accurate and sensible information), corresponding to soundness in our study; their intentions (e.g. to help the public affected by a crisis), corresponding to our helpfulness theme; and their adherence to a set of accepted principles, e.g. as imposed by the profession on emergency managers, who were the source of our messages. Furthermore, trust models discuss the trustor's propensity to trust (Mayer et al., 1995), which, in our study, seemed to be mainly determined by previous experience of flooding.

A final interesting finding from this study was the demonstration that greater comprehension is associated with greater trust. This finding provides empirical evidence of the role that clear crisis communications—through plain language and/or translation—can play in establishing a relationship of trust between emergency managers and the public, thus leading to higher compliance with instructions for crisis preparedness. A similar result, although related to advertisement disclaimers, is reported in Herbst et al. (2013).

This study has several limitations, particularly the high level of English proficiency of most participants, and the limited number of messages that were evaluated. Larger-scale experimental studies with different setups are warranted. Additional research should focus on: the impact of different PE levels; the impact of labelling human translations as post-edits; different language pairs; and end users less familiar with MT (Figure 2). It might also be interesting to observe end users' interactions with MT and to explain them using a trust and credibility lens (see e.g. Gao et al. 2014). Finally, future qualitative research could help determine the directionality of the relationship between comprehension and trust.

Acknowledgement This research has been funded by Science Foundation Ireland through the SFI Research Centres Programme and is co-funded under the European Regional Development Fund through Grant n. 13/RC/2106. Data from this study are available upon request.

References

Alexander, David. 2002. *Principles of Emergency Planning and Management.* Dunedin Academic Press Ltd, Edinburgh, United Kingdom.

Alexander, David, and Gianluca Pescaroli. 2019. The Role of Translators and Interpreters in Cascading Crises and Disasters: Towards a Framework for Confronting the Challenges. *Disaster Prevention and Management*, 29(1). doi:10.1108/DPM-12-2018-0382.

Anderson, John, Lorinda Mak, Aram Keyvani Chahi, and Ellen Bialystok. 2018. The Language and Social Background Questionnaire: Assessing Degree of Bilingualism in a Diverse Population. *Behavior Research Methods*, 50(1):250–263.

Braun, Virginia, and Victoria Clarke. 2012. Thematic Analysis. *APA Handbook of Research Methods in Psychology: Vol. 2: Research Designs*. American Psychological Association, Washington, 57–71.

Briggs, Neil. 2018. Neural Machine Translation Tools in the Language Learning Classroom: Students' Use, Perceptions, and Analyses. *JALT CALL Journal*, 14(1):2–24.

Cadwell, Patrick. 2015. *Translation and Trust: A Case Study of how Translation Was Experienced by Foreign Nationals Resident in Japan for the 2011 Great East Japan Earthquake*. PhD thesis, Dublin City University.

Cadwell, Patrick, Sharon O'Brien, and Carlos Teixeira. 2017. Resistance and Accommodation: Factors for the (Non-) Adoption of Machine Translation among Professional Translators. *Perspectives*, 26(3):301–321.

Cadwell, Patrick, Sharon O'Brien, and Eric DeLuca. 2019. More than Tweets: A Critical Reflection on Developing and Testing Crisis Machine Translation Technology. *Translation Spaces*, 8(2):300–333.

Castilho, Sheila, and Sharon O'Brien. 2016. Evaluating the Impact of Light Post-Editing on Usability. *10th International Conference on Language Resources and Evaluation*, Portorož, Slovenia, 310–316.

Central Statistics Office. 2016. *Census 2016 – Non-Irish Nationalities Living in Ireland.* https://bit.ly/2vNMCsv (Accessed 3 February 2020).

Depraetere, Ilse. 2010. What Counts as Useful Advice in a University Post-Editing Training Context? Report on a Case Study. *14th Annual Meeting of the European Association for Machine Translation*, Saint-Raphaël, France.

Federici, Federico, Brian Gerber, Sharon O'Brien, and Patrick Cadwell. 2019. *The International Humanitarian Sector and Language Translation in Crisis Situations. Assessment of Current Practices and Future Needs*. INTERACT The International Network on Crisis Translation, London, Dublin, Phoenix.

Gao, Ge, Bin Xu, Dan Cosley, and Susan Fussell. 2014. How Beliefs about the Presence of Machine Translation Impact Multilingual Collaborations. *17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Baltimore, USA, 1549–1560.

Gaspari, Federico. 2007. *The Role of Online MT in Webpage Translation.* PhD thesis, The University of Manchester.

Guerberof, Ana. 2013. What do Professional Translators Think about Post-Editing? *The Journal of Specialised Translation*, 19:75–95.

Herbst, Kenneth, Sean Hannah, and David Allan. 2013. Advertisement Disclaimer Speed and Corporate Social Responsibility: "Costs" to Consumer Comprehension and Effects on Brand Trust and Purchase Intention. *Journal of Business Ethics*, 117: 297–311.

Jeffers, James. 2011. The Cork City Flood of November 2009: Lessons for Flood Risk Management and Climate Change Adaptation at the Urban Scale. *Irish Geography*, 44(1):61–80.

Koponen, Maarit. 2015. How to Teach Machine Translation Post-Editing? Experiences from a Post-Editing Course. *4th Workshop on Post-Editing*

Technology and Practice, Miami, United States, 2–15.

Lewis, William. 2010. Haitian Creole: How to Build and Ship an MT Engine from Scratch in 4 days, 17 hours, & 30 minutes. *14th Annual Meeting of the European Association for Machine Translation*, Saint-Raphaël, France.

Liu, Rui, Runtong Zhang, and Xinyi Lu. 2018. An Empirical Study on the Relationship between the Satisfaction of Internet Health Information and Patient Compliance Based on Trust Perspective. *8th International Conference on Information Communication and Management*, Edinburgh, United Kingdom.

Major Emergency Management. 2016. *A Framework for Major Emergency Management*. https://bit.ly/2Vxqb3G (Accessed 3 February 2020).

Mayer, Roger, James Davis, and David Schoorman. 1995. An Integrative Model of Organizational Trust. *Academy of Management Review*, 20(3): 709–734.

Moorkens, Joss, and Sharon O'Brien. 2013. User Attitudes to Post-Editing Interface. *MT Summit XIV Workshop on Post-editing Technology and Practice*, Nice, France, 19–25.

Nurminen, Mary, and Niko Papula. 2018. Gist MT Users: A Snapshot of the Use and Users of One Online MT Tool. *21st Annual Conference of the European Association for Machine Translation*, Alacant, Spain, 199–208.

O'Brien, Sharon, and Federico Federici. 2019. Crisis Translation: Considering Language Needs in Multilingual Disaster Settings. *Disaster Prevention and Management*, 29(1). doi:10.1108/DPM-11-2018-0373.

O'Brien, Sharon, Federico Federici, Patrick Cadwell, Jay Marlowe, and Brian Gerber. 2018. Language Translation during Disaster: A Comparative Analysis of Five National Approaches. *International Journal of Disaster Risk Reduction*, 31:627–636.

O'Brien, Sharon, and Patrick Cadwell. 2017. Translation Facilitates Comprehension of Health-Related Crisis Information: Kenya as an Example. *Journal of Specialised Translation*, 28:23–51.

Porro Rodríguez, Victoria, Lucia Morado Vázquez, and Pierrette Bouillon. Study on the Use of Machine Translation and Post-Editing in Swiss-Based Language Service Providers. *Parallèles*, 29(2):19–35.

Santos-Hernández, Jenniffer, and Betty Hearn Morrow. 2013. Language and Literacy. Deborah S. K. Thomas, Brenda D. Phillips, William E. Lovekamp, and Alice Fothergill (eds). *Social Vulnerability to Disasters*, 2nd ed. CRC Press, Florida, 265–280.

Screen, Benjamin. 2019. What Effect does Post-Editing Have on the Translation Product from an End-User's Perspective? *The Journal of Specialised Translation*, 31:133–157.

Sherly, Mazhuvanchery, Subhankar Karmakarm, Devanathan Parthasarathy, Terence Chan, and Christian Rau. 2015. Disaster Vulnerability Mapping for a Densely Populated Coastal Urban Area: An Application to Mumbai, India. *Annals of the American Association of Geographers*, 105(6). doi:10.1080/00045608.2015.1072792.

Soricut, Radu, and Abdessamad Echihabi. 2010. TrustRank: Inducing Trust in Automatic Translations via Ranking. *48th Annual Meeting of the Association for Computational Linguistics*, Uppsala, Sweden, 612–621.

Specia, Lucia, and Kashif Shah. 2018. Machine Translation Quality Estimation: Applications and Future Perspectives. *Translation Quality Assessment: From Principles to Practice*. Springer, Cham, Switzerland.

TAUS 2010. *MT Post-Editing Guidelines.* https://bit.ly/2S15RH6 (Accessed 23 April 2020).

Teixeira, Carlos. 2014. Perceived vs. Measured Performance in the Post-Editing of Suggestions from Machine Translation and Translation Memories. *Third Workshop on Post-Editing Technology and Practice*, Vancouver, Canada, 45–59.

Toral, Antonio, Sheila Castilho, Ke Hu, and Andy Way. 2018. Attaining the Unattainable? Reassessing Claims of Human Parity in Neural Machine Translation. *Third Conference on Machine Translation*, Brussels, Belgium, 113–123.

Tourism Ireland 2018. *Visitor Facts and Figures.* https://bit.ly/36X51Qf (Accessed 3 February 2020).

Translators without Borders 2016. *Translators without Borders Develops the World's First Crisis-Specific Machine Translation System for Kurdish Languages.* https://bit.ly/3bjSfiq (Accessed 3 February 2020).

Translators without Borders 2019. *Misunderstanding + Misinformation = Mistrust.* https://bit.ly/37PRtHH (Accessed 3 February 2020).

Van Egdom, Gys-Walt, and Mark Pluymaekers. 2019. Why Go the Extra Mile? How Different Degrees of Post-Editing Affect Perceptions of Texts, Senders and Products among End Users. *Journal of Specialised Translation*, 31:158–176.