Applying Machine Translation to Psychology: Automatic Translation of Personality Adjectives

Ritsuko Iwai[1,2], Daisuke Kawahara[1], Takatsune Kumada[1,2], Sadao Kurohashi[1]

[1]Kyoto University / Yoshida Honmachi, Sakyo-ku, Kyoto-shi, 606-8501, Kyoto, JAPAN

[2]RIKEN BSI-TOYOTA Collaboration Center / Hirosawa 1-1, Wako-shi, Saitama, 351-0198, JAPAN

{ritsuko.iwai,dk,t.kumada,kuro}@i.kyoto-u.ac.jp

Abstract

We introduce our approach to apply machine translation to psychology, especially to translate English adjectives in a psychological personality questionnaire. We first extend seed English personality adjectives with a word2vec model trained with web sentences, and then feed the acquired words to a phrase-based machine translation model. We use Moses trained with bilingual corpora that consist of TED subtitles, movie' subtitles and Wikipedia. We collect Japanese translations whose translation probabilities are higher than .01 and filter them based on human evaluations. This resulted in 507 Japanese personality descriptors. We conducted a web-survey (*N*=17,751) and finalized a personality questionnaire. Statistical analyses supported the five-factor structure, reliability and criterion-validity of the newly developed questionnaire. This shows the potential applicability of machine translation to psychology. We discuss further issues related to machine translation application to psychology.

1 Introduction

This study introduces an example of the application of machine translation (MT) to psychology for academic research purposes. Translation is a critical part in psychological studies using questionnaires. Developing psychologically equivalent questionnaires across languages and cultures involves careful consideration and requires good knowledge of both the source and target languages and familiarity in psychological theories and concepts. The construction process requires conceptual equivalence as well as semantic equivalence (Herdman et al., 1998). Considering a typical phrase in an English questionnaire, "he is open to experiences," for example, it is not difficult to understand with a proficiency in English, but it is a difficult task to translate this description of personality in simple and intuitively understandable words (semantic equivalence), because "open" in Japanese is not used with abstract words. In addition, the translated items must reflect similar psychological concepts among Japanese respondents (conceptual equivalence).

The lists of English personality adjectives are generally difficult to translate because of their semantic ambiguities. For example, a typical personality adjective, "complex," has four meanings in the online Cambridge English dictionary[1] : Having a lot of different but related parts, being difficult to understand because of relatedness of parts, building, and bad feeling. In addition, multiple translations are listed in the Weblio English-Japanese dictionary[2] for each meaning. In the Japanese version of the Ten Item Personality Inventory (TIPI-J; Oshio et al., 2012), "complex" is translated as "変わった考えを持つ"/having unique ideas. It is difficult to judge to what extent the translation reflects on "complex." Furthermore, "変わった" has nuances of odd and strange in Japanese.

Therefore, it is often difficult to find adequate words that satisfy both conceptual and semantic equivalence. To resolve the issue, we use word embeddings and phrase-based statistical machine translation to translate English personality adjectives into Japanese. We regard a bag of personali-

---

[1] https://dictionary.cambridge.org
[2] https://ejje.weblio.jp

ty related words as covering the concepts. The translated results should reflect the equivalent concepts and semantics.

2 Related Work

2.1 Big-Five

Big-Five is one of the most widely used frameworks to understand human universal personality (e.g., McCrae and Costa, 1997). It assumes that human individual differences in personality are describable in five broad traits; Extraversion (extraverted, sociable, and talkative), Agreeableness (cooperative, agreeable, and warm), Conscientiousness (self-disciplined, well-prepared, and self-motivated), Neuroticism (anxious and emotional), and Openness-to-Experiences (analytic, creative, and curious)[3].

The framework comes from the lexical approach. Researchers collected adjectives that describe human personality from dictionaries, repeated human evaluations and identified the five-factor structure (e.g., Goldberg, 1992; Norman, 1963).

2.2 Translation of Big-Five Questionnaires into Japanese

For Japanese, two previous studies were devoted to obtaining adequate translations of personality adjectives. Wada (1996) listed all the translations of the 300 English personality-like adjectives in Adjective Checklist (Gough and Heilbrun, 1983), referring to an English-Japanese dictionary. She constructed the Big Five Scale (BFS) with 60 items evaluated by university students. Another study was conducted by Oshio et al. (2012). They translated the Ten Item Personality Inventory and evaluated the items five times by means of respondent surveys (TIPI, Gosling et al., 2003). They also asked the original authors if the backward translation appropriately reflected the five original personality concepts. The final Japanese version of the TIPI (TIPI-J) used relatively long and explanatory phrases as translations of simple English adjectives.

2.3 Automatic Translation of Personality Adjectives

Ueda et al. (2016) introduced an approach to acquire the Japanese translations of English personality adjectives. They used 20 personality adjectives derived from TIPI (Gosling et al., 2003) and acquired words related to these adjectives, using a word2vec model trained with 50 million web-sentences. Having personality adjectives in a bilingual corpus, they searched for bilingual corpora and combined the three bilingual corpora, 0.2 million TED subtitles, 1.2 million movie subtitles, and 0.4 million Wikipedia sentences. Iwai et al. (2017) selected 109 candidates from the list by Ueda et al. (2016) and conducted a web-survey with 500 young and 500 older adults. They identified the five-factor structure in both young and older samples. Although they planned to use four words per trait, they found only three words for EX and AG, and two words for CO. The numbers were unsatisfactory to construct a personality questionnaire. In this study, we improve the procedure to acquire the translation candidates and finalize the 20-item personality questionnaire.

3 Automatic Translation of Personality Adjectives

In this section, we introduce our method to translate psychological questionnaires.

3.1 Preparation for Bilingual Corpus

As a previous study (Ueda et al., 2016), to limit the scope to daily life contexts, we combine the bilingual corpora of TED talks, movie subtitles, and Wikipedia. The size is, however, slightly larger than Ueda et al. (2016), 2.1 million sentences, 0.3 million sentences more; TED[4] (0.2 million sentences), movie subtitles[5] (1.5 million sentences), and Wikipedia (0.4 million sentences, Chu et al., 2016).

3.2 Acquisition of Words Similar to Personality Adjectives.

In addition to TIPI (Gosling et al., 2013), we prepare the personality adjectives derived from Goldberg (1992) (Table 1). We feed the English personality adjectives to the word2vec[6] model trained with 50 million English web-sentences. The adjectives are, however, abstract and polysemous, and not necessarily limited to describe personality. We, thus, combine from two to four words to generate the averaged vectors within the same trait, in addition to feeding one adjective to

---

[3] In the following sections, Extraversion is abbreviated as EX, Agreeableness as AG, Conscientiousness as CO, Neuroticism as NE, and Openness-to-Experiences as OP.

[4] https://wit3.fbk.eu
[5] http://diates.lingfil.uu.se
[6] https://code.google.com/p/word2vec

Table 1 *Examples of Personality Adjectives, Related Words, and Phrase-Based Translations*

Trait	+/-	English	Sim.	Japanese	Prob.	Method
EX		courageous	.682	勇敢だ*	.364	combination
	+	playful	.695	ふざける	.333	combination
		talkative	1.000	話し好きだ	.308	combination
		disconcert	.546	当惑*	1.000	combination
	-	timid	1.000	臆病だ	.375	individual
		timid	1.000	内気だ	.250	individual
AG		respectful	.690	尊敬	.333	combination
	+	merciful	.290	慈悲深い	.235	individual
		dignify	.667	威厳	.200	combination
		arrogant	.731	傲慢だ	.235	combination
	-	cynical	.764	皮肉だ	.231	combination
		selfish	1.000	利己的だ	.229	individual
CO		honest	.657	正直だ	.274	combination
	+	trustworthy	.696	信頼できる	.254	combination
		conscientious	1.000	誠実だ	.308	combination
		insensitive	.712	無神経だ	.500	combination
	-	insensitive	.712	鈍感だ	.500	combination
		foolish	.731	愚かだ	.345	combination
NE		unhappy	.755	不幸だ	.234	combination
	+	apprehensive	.747	うろたえる	.200	combination
		emotional	1.000	感情*	.199	combination
		restful	.694	安らかだ	1.000	combination
	-	carefree	.668	のんきだ	.345	combination
		tolerable	.731	我慢	.345	combination
OP		inventive	.736	独創的だ*	.143	combination
	+	creative	1.000	創造的だ*	.115	combination
		intelligent	1.000	知的だ*	.111	combination
		unsophisticated	1.000	浅い	.250	individual
	-	vulgar	.728	下品だ	.222	combination
		simplistic	.783	単純だ	.158	combination

Note: + = positive; - = negative; Sim. = similarity values of word embeddings; * = the translations that are in Iwai et al. (2017); Prob.=translation probabilities; Sim. 1 = the words in the previous studies (Goldberg, 1992; Gosling et al., 2003); Individual=single word; Combination=combined vectors.

the model and list the related words when cosine similarities are higher than .2. In the case of related words with multiple factors, each word is associated with the trait that has the highest similarity. "Arrogant", for example, is similar to both AG minus and CO minus. In this case, the similarity to AG minus is higher (.783) than CO minus (.354), as such, we associate "arrogant" with AG minus.

As a result, we acquired a total of 200 unique candidate words in total. The list includes the words that are not in either Goldberg (1992) or Gosling et al. (2003), such as courageous, playful, and thoughtful.

3.3 Automatic Translation

Referring to Ueda et al. (2016), we also use the phrase-based statistical machine translation system, Moses[7]. Table 1 indicates examples of translation results. Using the bilingual corpora in section 3.1, we developed a phrase table and extracted 728 unique translations of the personality adjectives-related words in the table with more than a .01 translation probability.

Filtering: We filter translations and merge variant expressions by using a Japanese morpheme analyzer JUMAN++[8] and a Japanese dependency and case structure analyzer KNP[9]. For filtering, we refer to parts of speech and information in JUMAN++ dictionaries and collect the translations under the conditions of content words, either adjectives, verbs or nouns and are not humans, because conjunctions and dummy nouns are also in the table. Furthermore, this procedure makes it possible to merge expression varieties such as 落ち着く and 落ち着いている into 落ち着く. The procedures result in 631 translations.

Human evaluations: In spite of filtering, the automatic filtering fails to exclude translations unrelated to personality. Such translation errors include mainly those presumingly caused by alignment errors. The errors, for example, are 綿花/cotton as the translation of *"indifferent,"* 耳/ ear for *"stubborn,"* and 訪米/visit to the United States for *"anxious"*. As a result, we regard 507 translations as personality descriptors (80.3 %).

Comparisons with previous studies: Merging the 109 personality descriptors in Iwai et al. (2017) and the above 507 ones results in a total of 559 unique personality descriptors. A total of 52 personality descriptors are unique in Iwai et al. (2017) and 450 are unique in the current list while 57 personality descriptors are in common. Newly-acquired translations include 積極/ "active" and 陽気だ/ "cheerful" while おとなしい/ "silent" and 内向的だ/ "introverted," for example, do not appear at this time.

The personality descriptor rate is similar to the previous study by Iwai et al. (2017) (81.0%). However, we acquire more varieties of personality descriptors that were not enough in the previous studies.

4 Development and Evaluations of a Japanese Personality Questionnaire

To select the twenty-items, we conducted a web-survey and statistical analysis to identify the five-factor structure, calculate reliability, and examine validity. Exploratory Factor Analysis (EFA) is a statistical approach to extract common factors across measured variables based on correlation coefficients (Fabrigar et al., 1999). In constructing a psychological questionnaire, it is important to evaluate reliability and validity. Reliability indicates how responses are reliably produced. Internal consistency assumes that a person tends to similarly answer items within the same trait, which Cronbach's α indicates (Cronbach, 1951). Furthermore, a psychological questionnaire must measure the targeted concepts, which is named as validity. One method to assess validity is criterion-validity. Criterion-validity investigates correlations between the latent variables in the newly constructed questionnaire and the corresponding latent variables in a "criterion" questionnaire. The correlations are expected to be high between similar latent variables and low between unrelated latent variables. We use TIPI-J (Oshio et al., 2012) for this validity evaluation.

4.1 Web-Survey

Participants: We conducted a web-survey on registrants of the Human Information Database by NTT Data Institute of Management Consulting, Inc in October 2016 (*N*=17,751, *Female*=11,037, *Mean*=49.8 years old, *SD*=13.4).

Tested Items: To conduct a web-survey, two psychologists (the first and the third authors) selected and modified the translations into questionnaire-item formats. We had limited resources and only 51 items were testable. In addition to the 18 items from Iwai et al. (2017), we, thus,

---

reviewed the list carefully and selected 33 items acquired from the list in Section 3.3. We basically attempted to choose words that are not in Iwai et al. (2017) and from EX plus (e.g., 陽気だ/ "cheerful" and 積極/ "active"), AG minus (e.g., 横柄だ/ "arrogant" and 傲慢だ/ "arrogant"), CO plus (e.g., 公正だ/ "fair" and 正直だ/ "honest"), NE minus (e.g., 安らかだ/ "peaceful" and のんきだ/ "carefree"), and OP minus (e.g., 単純だ/ "simplistic" and 浅い/ "unsophisticated").

OP minus translations are especially difficult. The OP minus adjectives mainly consist of the adjectives with an "un" prefix. As a result, we have a very limited number of candidate words in OP minus. "浅い, "an example of translation of "unsophistication," literally means shallow, is modified into 興味が浅い/ "have an shallow interest."

Participants completed the 51 items, using a seven-point scale from 1= strongly disagree to 7 =strongly agree.

Ten Item Personality Inventory-Japanese (TIPI-J): Participants also completed another Big-Five questionnaire, TIPI-J (Gosling et al., 2003; Oshio et al., 2012) using a seven-point scale (EX: *M*= 7.4, *SD*=2.3, α^{10}=.49; AG: *M*= 9.5, *SD*=2.0, α=.37; CO: *M*= 8.2, *SD*= 2.2, α=.51; NE: *M*= 7.9, *SD*=2.1, α=.46; OP *M*= 8.0, *SD*=2.1, α=.39).

4.2 Results

Selection of twenty-items: We repeated the exploratory factor analysis (EFA) using a maximum likelihood with robust standard errors method with a geomin rotation[11] to select the twenty items (four items for each factor). Based on the first EFA with the 51 items, we eliminated the 9 items that loaded highly on multiple factors. Next, we conducted the second EFA with the 41 items, reviewed the loading matrix, and selected four items for each trait based on the following standards: An item (1) which loading was over |.50| and (2) the secondary loading did not exceed |.35|. In the cases that a trait failed to acquire four items in reference on the standards, we selected the items by loading. Selecting AG items, however, was not straightforward. Although "安らか/peaceful" indicated high loading in AG, the word caused confusion because it is often used to describe facial expressions or

[10] α= Cronbach α
[11] For details about EFA rotations, please refer to Browne (2001).

Table 2 *The Final 20-Items*

Traits	Items
EX	陽気だ (N), 活発だ(C), 積極 (N), 話し好きだ (N)
AG	穏やかだ (N), 協力 (C), 温和だ (O), 同調 (O)
CO	無神経だ (N), 鈍感だ (N), 無責任だ (C), いい加減だ (N)
NE	心配 (C), 自責 (O), 混乱 (C), 感情 (C)
OP	分析 (C), 考察 (O), 独創 (C), 知性 (O),

Note: O = the words that appear only in Iwai et al. (2017); N = the words that appear only in this study; C = the words in common between Iwai et al. (2017) and the current study.
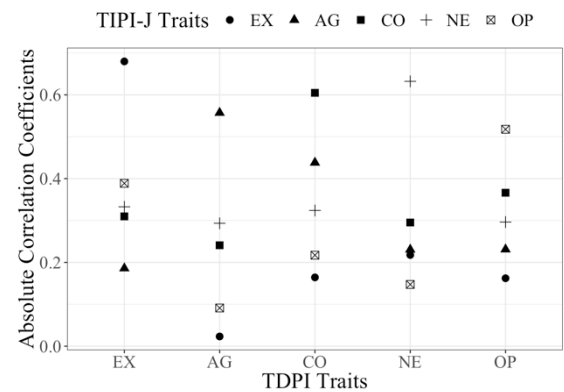

Figure 1 Correlation coefficients between Trait Descriptors Personality Inventory (TDPI) and Ten Item Personality Inventory-Japanese (TIPI-J).

mood, rather than personality. "協力的 /cooperative" seemed more acceptable among the other candidates in terms of semantics. We, thus, conducted the two patterns of EFA using the 20 items. The model fit indices were slightly better in the "安らか/peaceful" version but had very small differences (CFI[12] = .978 vs. .972, TLI[13]=.958 vs. 947, RMSEA[14]=.029 vs. .033)[15] and the factor loading patterns were similar. We, thus, decided to finalize the 20-items including "協力的." Table 2 indicates the twenty-words in the items. We name the questionnaire as Trait Descriptors Personality Inventory (TDPI).

Descriptive Statistics and Reliability: We calculated descriptive statistics for each trait and internal consistency (EX: *M*= 15.9, *SD*=4.2,

[12] Comparative Fit Index
[13] Tucker Lewis Index
[14] Root Mean Square Error of Approximation
[15] The model fit indices are considered as excellent when CFI and TLI > .950, RMSEA < .03 and good when CFI and TLI > .900, RMSEA < .05 (Marsh et al., 2009).

α=.82;AG: *M*= 17.8, *SD*=2.0, α=.78;CO: *M*= 17.9, *SD*= 4.2, α=.79;NE: *M*= 15.9, *SD*=4.2, α=.66;OP *M*= 17.8, *SD*=3.4, α=.74). Cronbach's αs were substantial. The means of correlation coefficients between the traits were quite low (the mean *r*=.22), which indicates that each trait was differentiated from other traits.

Criterion-validity. We calculated correlation coefficients of trait scores in TDPI with those in TIPI-J (Figure 1). The means of correlation coefficients between the same traits were high (*r* =.58 ~.68). On the other hand, the means of not-corresponding correlation coefficients were low (*r*=.25). The results indicate that the two questionnaires measure similar psychological factors and differentiate the similar factors from the factors that are hypothesized as different ones.

5 Discussion

The results of the web-survey show applicability of MT to psychological studies, i.e., using MT to extracting candidate entries. However, we find three issues which need to be considered.

5.1 Limited Resources

The translation probability has a limited role in the procedure. Overall, translation probabilities are relatively low. Only 138 of 631 translations indicate more than a .1 translation probability. Furthermore, most of the errors are alignment errors. Such errors imply that the bilingual corpus does not include enough translations of personality descriptors. The present study focuses on personality. However, there are many psychological questionnaires with English adjectives such as values (Schwartz et al., 2003) and interpersonal relationships (Fletcher et al., 2003). We expect that such contextual matters and not enough resources are the shared issue for those who are interested in using MT in practical usage.

5.2 Replicability and Stability

While this study indicates applicability to questionnaire development in psychology, it entails the issue of replicability and stability due to choices of a mono-lingual corpus for word embeddings. Our study uses the same size corpus for word embeddings as Iwai et al. (2017) did. However, 47.2 % of the translations are not replicated in the current study. Out of the final 20 items, the 5 items are in the previous study, the 7 items only from this study, and 8 items are in common.

Inconsistent replicability is not due to translation. Our procedure is phrase-based and we use the extended bilingual corpus of Iwai et al. (2017). Even if we ignore translation probabilities and review all the results, many of them are not in the list. This indicates that different personality adjectives-related words were fed into Moses, because all the phrases in English are aligned to the particular phrases in Japanese. However, the translations peculiar in the previous studies are also good as personality descriptors and some of them remain as the final items.

As the current study suggests, it is better to acquire substantial candidates with limited resources as in Section 5.1 and it is better to repeat the procedures.

5.3 Expert Knowledge

The two previous issues are all solved by using the psychologists' expert knowledge. The previous studies (Iwai et al., 2017; Ueda et al., 2016) and this study demonstrate that word embeddings and MT allow researchers to collect personality-related English words and Japanese translation candidates and such candidates are tolerable to use as psychological items with expert knowledge. On the other hand, it is still just at the beginning of the step to implement MT into psychological studies. It is highly appreciated that the manual parts are reduced and replaced with technical improvements in NLP and MT.

6 Conclusions

MT allowed us to collect candidates of Japanese personality descriptors. We manage to construct a new personality questionnaire that consists of only MT-extracted words. To the best of our knowledge, this is the only personality measurement developed using Natural Language Processing (NLP) techniques such as word embeddings and phrase-based statistical MT. The questionnaire is practically usable in psychological studies. The study provides evidence to extend applicability of MT to another research field. On the other hand, the expert knowledge is critical, at least, in the target language and culture, to design a questionnaire and items. Such experts' efforts are expected to be reduced with more adequate parallel corpora and further examination to justify word embeddings.

References

Michael W. Browne. 2001. An Overview of Analytic Rotation in Exploratory Factor Analysis. *Multivariate Behavioral Research*, 36(1):111–150.

Chenhui Chu, Toshiaki Nakazawa, and Sadao Kurohashi. 2016. Integrated Parallel Sentence and Fragment Extraction from Comparable Corpora. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 15(2):10–22.

Lee J. Cronbach. 1951. Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 16(3):297–334.

Leandre R. Fabrigar, Duane T. Wegener, Erin J. Strahan, and Robert C. MacCallum. 1999. Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Assessment*, 4(3):272–299.

Garth J. O. Fletcher, Jeffry A. Simpson, and Geoff Thomas. 2000. The Measurement of Perceived Relationship Quality Components: A Confirmatory Factor Analytic Approach. *Personality and Social Psychology Bulletin*, 26(3):340–354.

Lewis R. Goldberg. 1992. The Development of Markers for the Big-Five Factor Structure. *Psychological Assessment*, 4(1):26–42.

Samuel D. Gosling, Peter J. Rentfrow, and William B. Swann Jr. 2003. A Very Brief Measure of the Big-Five Personality Domains. *Journal of Research in Personality*, 37(6):504–528.

Harrison G. Gough and Alfred B. Heilbrun. 1983. *The Adjective Check List Manual.* Consulting Psychologists Press, Palo Alto, 1983 edition.

Michael Herdman, Julia Fox-Rushby, and Xavier Badia. 1998. A Model of Equivalence in the Cultural Adaptation of HRQoL Instruments: The Universalist Approach. *Quality of Life Research*, 7(4):323–335.

Ritsuko Iwai, Takatsune Kumada, Daisuke Kawahara, and Sadao Kurohashi. 2017. Translating Big-Five Personality Constructs from English to Japanese, Using Statistical Machine Translation. Poster session presented at *the18th Annual Meeting of Society for Personality and Social Psychology,* San Antonio, U.S.A.

Herbert W. Marsh, Bengt Muthén, Tihomir Asparouhov, Oliver Lüdtke, Alexander Robitzsch, Alexandre J. S. Morin, and Ulrich Trautwein. 2009. Exploratory Structural Equation Modeling, Integrating CFA and EFA: Application to Students' Evaluations of University Teaching. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3):439–476.

Robert R. McCrae and Paul T. Jr. Costa. 1997. Personality Trait Structure as a Human Universal. *American Psychologist*, 52(5):509–516.

Atsushi Oshio, Shingo Abe, and Pino Cutrone. 2012. Development, Reliability, and Validity of the Japanese Version of Ten Item Personality Inventory (TIPI-J). *The Japanese Journal of Personality*, 21(1):40–52.

Shalom H. Schwartz, Gila Melech, Arielle Lehmann, Steven Burgess, Mari Harris, and Vicki Owens. 2001. Extending the Cross-Cultural Validity of the Theory of Basic Human Values with a Different Method of Measurement. *Journal of Cross-Cultural Psychology*, 32(5):519–542.

Shinpei Ueda, Daisuke Kawahara, Sadao Kurohashi, Ritsuko Iwai, and Takatsune Kumada. Automatic Translation of English Personality Adjectives into Japanese. In *Proceedings of 22nd Annual Meeting of Natural Language Processing,* Sendai, Japan 282–285.

Sayuri Wada. 1996. Construction of the Big Five Scales of Personality Trait Terms and Concurrent Validity with NPI. *The Japanese Journal of Psychology*, 67(1):61–67.