Developing Multilingual Questionnaires:
A Sociolinguistic Perspective[1]

Yuling Pan[1] and Marissa Fond[1]

[1]U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Abstract

Central to the development of multilingual questionnaires is a sound translation of the source questionnaire into target languages. This study aims to determine the types of translation issues that can lead to measurement errors in cross-cultural studies. Based on two multilingual projects conducted to cognitively test the translations of the 2010 U.S. Census questionnaire in four languages as well as the English original (Chinese, Korean, Russian, and Vietnamese) and the American Community Survey questionnaire in two languages (Chinese and Korean), we developed a coding scheme guided by sociolinguistic theories to systematically analyze translation issues. This paper discusses how the coding scheme can be useful in the development of multilingual questionnaires and suggests feasible solutions to translation issues, so as to ensure translation quality.

Key Words: survey translation, multilingual questionnaire, sociolinguistics, measurement errors

1. Introduction

Central to the development of multilingual questionnaires is a sound translation of the source questionnaire into target languages. Prior research on survey translation (e.g., Harkness et al., 2003; Pan & de la Puente, 2005; Forsyth et al., 2007) has a tendency to focus on the procedural aspect of the translation-review process, rather than the outcome. This paper demonstrates the challenges in evaluating translated survey questions and argues for the need to develop a translation-review framework guided by sociolinguistic theories to overcome these challenges. Based on findings from 112 cognitive interviews conducted by the U.S. Census Bureau to evaluate the 2010 Census questionnaire in five languages, we developed a coding scheme to classify translation challenges, other than translation errors, that are caused by different linguistic conventions (usage of words and grammar), cross-cultural communication norms (appropriate expressions of a concept), and social practices (knowledge needed to process a concept or to answer a question). We then applied the coding scheme in another study to evaluate the Chinese and Korean translations of the American Community Survey. Thus, this study aims to determine the types of translation issues that can lead to measurement errors in cross-cultural studies, and discusses how the coding scheme can be useful in the development of multilingual questionnaires. We also explore a number of feasible solutions to address the types of translation issues encountered, so as to ensure translation quality.

---

[1] *Disclaimer*: This report is released to inform interested parties of research and to encourage discussion. This views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

1.1 The Need for Translation Research

The development of multilingual questionnaires requires comprehensive, reliable translations from the source language into target languages. Researchers working on survey translation methods generally agree that an effective survey questionnaire translation is one in which the translated questionnaire "asks the same questions and offers the same response options as those provided in the source questionnaire" (Harkness & Schoua-Glusberg, 1998, p. 92). Achieving functional equivalence is deemed the goal of survey translation; however, achieving this goal can be deceptively complicated. Many questions remain, such as: How can functional equivalence be measured? How can the quality of translated material be consistently and reliably evaluated? Also, more importantly, once a translation has been through the expert review process, or quality control, how can any issues still causing difficulties be addressed?

The most common practice of survey translation follows a traditional approach in the process of translation. It usually consists of the following steps: translation, review, revision, and finalizing. This is a linear approach involving a single party at each step. Recent developments in survey translation research call for a more rigorous procedure in the translation review process and call for research in translation review methods. For example, Harkness et al. (2003) and Forsyth et al. (2007) proposed a "TRAPD" method, which stands for translation, review, adjudication, pretesting, and documentation. Also with the goal of establishing a more rigorous procedure, the U.S. Census Bureau established its own translation guidelines (de la Puente & Pan, 2004; Pan & de la Puente, 2005) and recommended a committee approach to translation and extensive pretesting of translations. These research efforts on translation evaluation methods signal a big step forward in the field because these approaches incorporate pretesting of translations with real respondents, and they involve multiple parties at each step in the process.

While this is a big step forward, this prior research places a heavy focus on the translation review procedure, rather than the final outcome. There is a lack of tools for the consistent evaluation of translation outcome (i.e., translation quality). Survey project managers and translation-reviewer committees in program areas still face the challenge of how to systematically assess the translation quality, particularly in a language that they do not speak or do not have expertise in. We believe there are four critical questions for us to consider: 1) How can we determine how successful a translation is? 2) When translation issues are identified in the expert review or cognitive testing process, how can we best articulate and describe the nature of the problems? 3) Can patterns of inadequacies in translated materials be identified so that they can be anticipated and addressed efficiently? 4) What are the feasible and effective solutions to the problems identified?

This study attempts to address these questions by developing a coding scheme that accounts for linguistic and sociocultural problems in survey translations. This coding scheme can be used as an assessment tool for translation evaluation.

1.2 The Sociolinguistics of Survey Translation

From a sociolinguistic perspective, a quality survey translation should be accurate and appropriate at three linguistic levels: lexical, syntactic, and pragmatic (Pan et al., 2007). The first level of appropriateness in translation is the lexical level, which is the smallest unit of a translation. This level concerns word choice and terminology. This means that the translation should consist of accurate, appropriate wording and use the correct terminology to convey the meaning of each individual word in the source text. The second level of appropriateness is at the syntactic level, which refers to the grammar of a

language. This means that the translation should follow the grammatical structure of the target language, including the syntactic rules and word order. It should be free of grammatical errors, and moreover it should sound natural in the target language. The third level is the pragmatic level, which refers to whether a translation can achieve its intended function, and whether respondents who read the translated questionnaire or other translated materials can understand the intended meaning and can take the required actions. This means that a translated text must fit within the sociocultural context in which the translated text is to be used. When we talk about translation issues at the pragmatic level, we talk about the frame of reference, or schema, that people rely on for interpreting a translated item. This involves sociocultural context, background knowledge or experience, and communication norms. Translation issues that are classified as pertinent to the pragmatic level of language, that is, how our target respondents interpret and react to the translated materials, are more subtle and difficult to overcome. They are usually not identifiable in the traditional translation-review process, however they are of critical importance to survey researchers because issues at the pragmatic level will hinder survey participation and will affect data quality. While traditional translation reviews address the lexical and syntactic levels, it is necessary to focus on the pragmatic level as well.

In order to ensure that translated questionnaires are appropriate at the pragmatic or "function" level as well as the lexical and syntactic level, we need to address three components in our analysis which we will call: Linguistic Forms, Cultural Norms, and Social Practices. A language encodes cultural values and salient social practices of a particular cultural group, and language use is always a reflection of cultural norms and social knowledge (Gumperz, 1999), and so these three components can serve as the guiding principles for us to evaluate the quality of a translation. First, we need to see if a translation is accurate in terms of Linguistic Forms, that is, language-specific rules for conveying meanings, including words, word order, and sentence structures. Secondly, we need to evaluate if a translation is appropriate in terms of whether it follows the Cultural Norms of communication, including appropriate communication style, culture-specific ways of showing politeness, and the use of the appropriate discourse structure for presenting information. Thirdly, we need to investigate whether there are differences in Social Practices between the source language's culture and the target culture, including differences in social institutions, educational systems, or respondents' experiences as influenced by culture and society. These differences can directly affect respondents' comprehension of the constructs being measured by survey questions. For example, if a concept does not exist in the target culture, it is very likely that the target language does not have a term to describe that concept. A good translation should find ways to successfully convey the concept or social practice that is foreign to the target populations.

2. The Coding Scheme

Based on these principles of sociolinguistics, we see the need to develop a coding scheme to evaluate the results of the translation process in a way that: 1) accounts for language use in social contexts; 2) describes the nature of problems found in translated texts; 3) can be quantified in order to uncover patterns or trends; 4) and, most importantly, offers solutions to address each type of problem.

To that end, we developed a coding scheme that includes three main categories that reflect the components of the lexical, syntactic, and pragmatic level of language, listed above: Linguistic Forms, Cultural Norms, and Social Practices. Two other categories,

User Errors and Translation Errors, are also used to make sure that all of the issues that are discovered are able to be classified appropriately. The source data for these categories are gathered by language experts who conduct the translation reviews and cognitive interviews. They have expertise in both language and culture, so they are quite attuned to the nuances in different translations, and they can recognize these issues when they arise.

In order to introduce the coding scheme before examining its development and use in two Census Bureau research projects, we briefly describe the categories here.

Codes	Explanations
Linguistic Forms (LF)	This category classifies issues in a translation that are due to vocabulary, grammar, usage conventions, etc.
Cultural Norms (CN)	This category refers to issues in which concepts that are expressed one way in English (the source language) are expressed in a different way in the target language (e.g. address conventions, numbering, kinship terms, conversational norms, etc.).
Social Practices (SP)	This category classifies concepts that can be described in English but cannot be translated into the target language because either the concept does not exist in that culture, or respondents have no experience with the concept.
Translation Errors (TE)	This category refers to translation problems that are simple mistakes (e.g. omissions, typos) that can be easily corrected.
User Errors (UE)	This code refers to actions taken by respondents while reading or answering a questionnaire that they themselves identify to be mistakes (easily corrected errors, for example those caused by inattention).

Table 1: The coding scheme.

Next we will describe the development and testing of the coding scheme in two multilingual projects undertaken by the U.S. Census Bureau.

3. Two Multilingual Projects

Two multilingual projects were conducted to develop and apply the coding scheme. In the first project, the 2010 Census questionnaire was pretested in five languages (English, Chinese, Korean, Russian, and Vietnamese) using cognitive interviewing techniques. Cognitive interviewing is a popular method for pretesting survey instruments and has been used to better understand survey respondents' interpretations of questions and the thought processes that they use to answer them (see Willis, 2005). The results from the cognitive testing (which consisted of the language experts' recommendations based on the cognitive testing they conducted) were instrumental in the development of the coding scheme. In the second project, the translation of the American Community Survey questionnaire (ACS) was reviewed numerous times by translation teams and then cognitively tested in two languages (Chinese and Korean); the results of this project were used to test the application of the coding scheme.

3.1 Results from the 2010 Census Questionnaire Project
The 2010 Census Questionnaire Project includes data from English, Chinese, Korean, Russian, and Vietnamese. The results reported here are based on the language experts'

reviews of the insights collected through the cognitive interview process; that is to say, the language expert created a final report summarizing the findings from the interviews in a given language, and provided his or her recommendations on each issue that was determined to be problematic. This final report was used to see if the issues discovered could be separated into groups, or if patterns were visible across issues or across languages. Based on this preliminary overview, the coding scheme was developed, and the issues discovered were evaluated and subsequently coded. The chart below (Figure 1) summarizes the findings of types of issues identified in the four target languages (excluding the English source material) during the cognitive testing process. There are four main types of issues coded: LF, CN, SP, and TE. We did not find any user error problems. This may be due to the fact that the 2010 Census questionnaire contains only ten demographic questions, and it is relatively simple in terms of form navigation.

Regarding these four types of issues, it quickly became apparent that LF and CN issues were the most common, as they comprised 33% and 39% of the total 163 issues discovered, respectively. The remaining issues were divided between SP and TE, with 18% and 10% respectively.
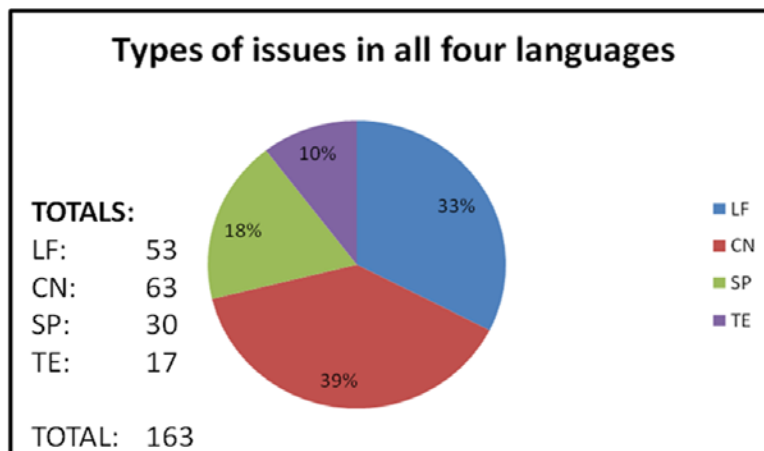
Figure 1: Types of issues in all four languages from the pretesting of the 2010 Census Questionnaire Project.

Next, in order to deepen our understanding of the issues and how the coding scheme works, we wanted to see if each of the four target languages seems to have the same pattern of issues identified. The chart below (Figure 2) shows that LF and CN issues constitute the majority in each language, as we already discussed. Korean and Vietnamese show similar proportions, but in Chinese, the LF category comprises a much larger percentage of total errors. We can use this information to focus our attention on explaining why that might be; LF might be much higher than it is in all the other languages because the Chinese translation had an unusually high number of issues with grammar; there were many examples of overly complex sentence structure (in particular, structures that were faithful to the English original rather than adjusted to typical Chinese structures) that were not present in the translations in the other languages.
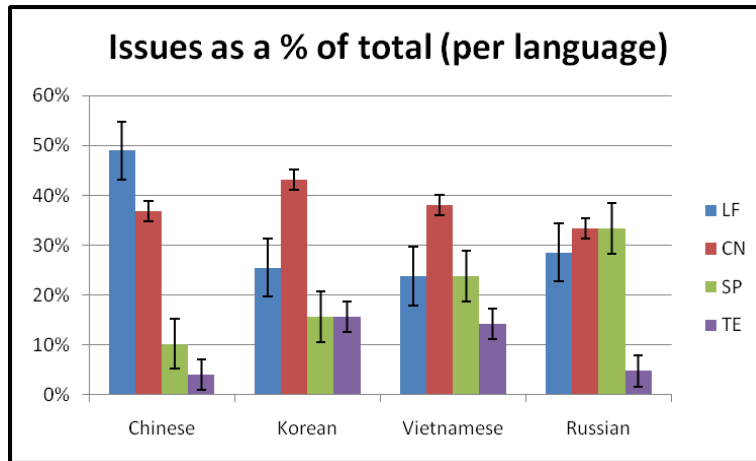
Issues as a % of total (per language)

Figure 2: Issues as a % of total (per language) in the 2010 Census Questionnaire Project.

It is important to emphasize that the coding scheme (as well as the results that it generates) constitutes a tool rather than an end in itself. In order to reach the goal of producing appropriate and accurate translations, it is necessary to understand the unique properties of each individual issue that arises in a given language, in a given questionnaire. So to describe the nature of the translation issues quantified in these results, and how they were coded, we can review some examples from the 2010 Census questionnaire translations, beginning with Question 2, seen in Figure 3. The examples are presented merely for the purpose of demonstrating how we coded the translation problems based on cognitive interview summaries from the language teams.

2. Were there any additional people staying here April 1, 2010 that you did not include in Question 1?
Mark ☒ all that apply.

☐ Children, such as newborn babies or foster children
☐ Relatives, such as adult children, cousins, or in-laws
☐ Nonrelatives, such as roommates or live-in baby sitters
☐ People staying here temporarily
☐ No additional people

Figure 3: Question 2 from the 2010 Census.

Regarding this question, there were translation issues in three of the four languages that were classified as related to Linguistic Forms. In Korean, "live-in baby sitter," seen here as part of the third response option, was translated as "always staying babysitter," which sounded as awkward and confusing in Korean as it does in English, indicating that the lexical items chosen to translate the term were inaccurate. In Chinese, the translation of the question was too close to the English syntactic structure; it did not correspond to the standard Chinese structure, and as a result it sounded awkward and was difficult for the Chinese-speaking respondents to process. In Vietnamese, the translation of the question included repetitive and unnecessary auxiliary verbs which obscured the meaning of the question; the verbs needed to be deleted to make the sentence sound natural in Vietnamese. These are all issues classified as Linguistic Forms issues because they implicate lexical items and syntax; the resulting awkwardness of the translations was due to the failure to adapt the English original to the linguistic features of the target language.

For another example, see Question 7 in Figure 4:

Figure 4: Question 7 from the 2010 Census.

An example of a Cultural Norms issue was present in the translation of this question into Korean. More than two thirds of Korean speakers had difficulty writing their age because the Korean convention of counting age is different from the American one. (In Korean culture, newborn babies are considered one year old, so someone who would be 50 years old as Americans count age would be considered 51 by Korean speakers.) In order to address this issue, the translation must specify that the form is asking about the American way of indicating age. This is a Cultural Norms issue because the misunderstanding arises not from the lexical items or the syntax of the translation, but from the cultural background of the speakers that causes them to interpret the question in a way that was not intended by the American survey designers.

Lastly, there were many examples of questions on the 2010 Census questionnaire that triggered issues with Social Practices. In Question 2, which has already been reviewed above in the context of issues with Linguistic Forms, the term "foster children" proved problematic for Chinese, Korean, and Vietnamese. There is an existing term for "foster children" in these languages, but the term means something quite different from the concept of "foster children" in the United States. In these languages, the term "foster children" was understood as "children under the temporary care of relatives or friends." The concept of a foster program administered and supported by the government was not retained in the translation. This type of program does not exist in China, Korea, or Vietnam, so there is no lexical item in the languages that can be used to describe it; a more descriptive phrase that provides additional information is necessary. Thus the recommendation to use a descriptive phrase "children in the foster program sponsored by the government" was accepted and incorporated in the final translations of the 2010 Census questionnaire.

Another example of a concept that is uniquely American is found in Question 3, seen here in Figure 5.

Figure 5: Question 3 from the 2010 Census.

Question 3 provides response options relating to the ownership of the residence, but the question itself proved to be difficult. Speakers of all four languages were unsure about what a "mobile home" was, as mobile homes are not common in other countries (and in fact, speakers of other languages such as Spanish will often adopt the American word "trailer" rather than describe the concept in their native language). A final example of a concept that could not be readily translated into another language clearly and unambiguously was "nursing home," found in the response options for Question 10 as seen in Figure 6:

Figure 6: Question 10 from the 2010 Census.

For the term "nursing home," the approximate translations were found to be inaccurate (or inadequate). The target languages have an equivalent term of "nursing home," but the meanings are different from the American concept. In Chinese, the translation meant, to various speakers, either a mental hospital (to Hong Kong Chinese) or a recreational resort (to Mainland Chinese); in Korean the translation meant a resting place; in Russian it meant a medical establishment; and in Vietnamese it meant a luxurious resort. Given that these cultures do not have experience with American-style nursing homes, the translations were unable to capture that meaning. A longer and more descriptive phrase is necessary in order to make sure that the individual filling out the questionnaire understands the question as it is intended. The recommendation to fix this problem is to have a descriptive phrase that conveys two meanings (a center for taking care of the elderly and for taking care of the very sick) in the English original instead of using the existing terminology.

Issues like the examples described here were analyzed in order to classify the underlying reasons why aspects of translations were found inadequate by the individuals reviewing them. The categories Linguistic Forms, Cultural Norms, and Social Practices were formed as a result, and applied in the second project, which was cognitive testing of the American Community Survey.

3.2 Results from the American Community Survey Project
The portion of the American Community Survey (ACS) examined in this project covers 21 housing-related questions and survey instructions, involving many complex questions and concepts that are not included on the 2010 Census questionnaire. For this survey, two translations were assessed: Chinese and Korean. And in contrast to the 2010 Census project, numerous rounds of translation review were conducted prior to testing the ACS questionnaire with the public. A total of 41 cognitive interviews were conducted with monolingual speakers of Chinese (19 interviews) and Korean (22 interviews). As a survey respondent fills out a questionnaire, a trained interviewer asks a series of probes designed to elicit the respondents' thoughts about what certain phrases mean, or what the motivation for asking a certain question is.

Because the coding scheme had already been developed based on the 2010 Census Questionnaire Project review, the results from these cognitive interviews could be coded accordingly. First, the cognitive interviewers, who are experts in cognitive interviewing and native speakers of the target language, attended training to learn the theoretical principles of the coding scheme and how to apply it. Then, they completed their interviews and wrote up summaries of the respondents' thoughts and opinions. They then coded translation issues that were uncovered during the interview as a result of the probes that were asked, and justified their application of each code. After they delivered their coded summaries to the Census Bureau, a Census Bureau research analyst read the

summaries and the justifications for each code, then confirmed each one. (The analyst also resolved any doubts that the cognitive interviewer may have had about which code to apply.) Therefore, for this project, at least two individuals evaluated each language issue using the coding scheme.

In the chart below (Figure 7), we see the distribution of the types of issues that were uncovered in the 41 cognitive interviews.
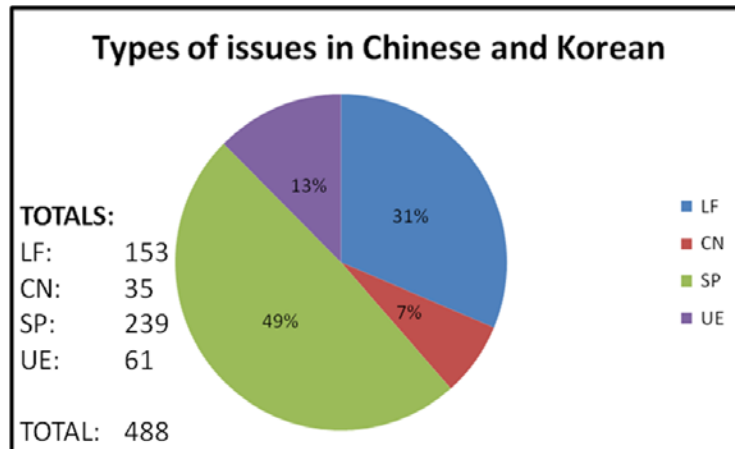

Figure 7: Types of issues in Chinese and Korean in the ACS questionnaire.

Notice that SP issues constitute the majority of the issues (49%) followed by the LF issues (31%) in the ACS study. This is most likely because, as stated earlier, the translations used in this study had undergone numerous rounds of translation review before the cognitive interviews were conducted. Social Practices issues are the most difficult to resolve through translation reviews, so it is not surprising that additional problems were found involving those issues. No translation errors were identified in the cognitive testing process. This is due to the fact that the translations were carefully reviewed prior to the cognitive testing. Small proportions of UE (13%) and CN issues (7%) were identified.

In the following chart (Figure 8), we can see the issues as a percentage of the total number of issues uncovered in each language.
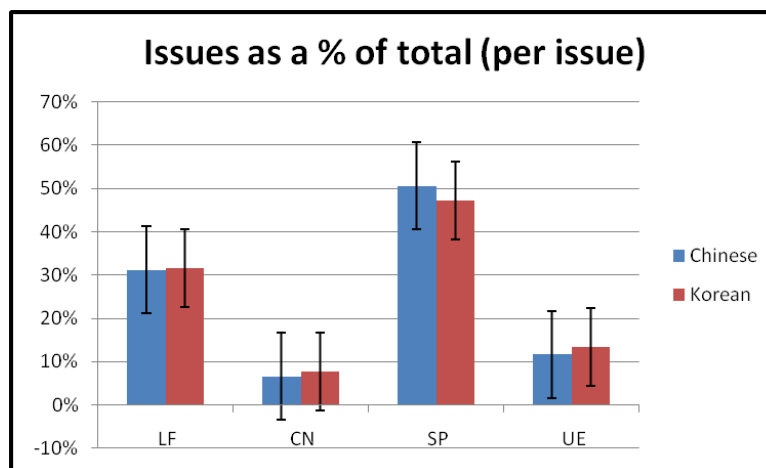
Figure 8: Issues as a % of total (per issue) in the ACS questionnaire.

In this chart, it is clear that the proportions of issues in Chinese and Korean are similar. This would be expected because the translations would have gone through expert review many times and the differences in quality among translation teams (which can be significant) would have been eliminated; also, most of the issues that remain are SP issues common to both Chinese and Korean societies (which do not have simple translation fixes) and residual LF issues (e.g. vocabulary that respondents interpret in ways that are different from how the translators did).

In order to illustrate how we applied the coding scheme, some examples of issues from the ACS follow.

The ACS Housing Question 1 asks: "Which best describes this building?" See Figure 9 below:

Figure 9: Question 1 from the ACS Housing.

After the respondent marked an answer, the interviewer administered a probe related to one of the response options: "What do you think this phrase 'Include all apartments, flats, etc., even if vacant' is saying? Does it sound natural in Chinese?" In an interview summary, the cognitive interviewer described a respondent's (R) interpretation as follows: "R mentioned that this phrase seems to mean to include all the rooms in R's house or apartment. R understood 单元房 ('flats') as 'rooms'." This was coded as LF because the meaning of the lexical item "flats" seems to be slightly different from what was intended; "flats" was meant to refer to apartments, made up of any number of rooms, but it was interpreted as "rooms." In this case, respondents could conclude that if their apartment included any vacant rooms (such as extra bedrooms, etc.), then they should be counted in the total.

The ACS Question 19a asks about home ownership; see Figure 10:

Figure 10: Question 19a from the ACS Housing.

The interview prompt was simple and open ended: "How did you come up with your answer? What do you think the question is asking?" For one interview, the interviewer summarized the problem: "R didn't understand the term 抵押贷款 (mortgage loan) as intended; she had a mortgage, but doubted that a mortgage was a 债务 (debt), which is a different concept in Chinese, among the general public." This indicates that although the technical term is 抵押贷款 (mortgage loan)，Chinese lay people are not aware that a mortgage means that the house is collateral and that it is a kind of debt. This is also indicates that 债务(debt) has a different socio-cultural meaning in Chinese from the American concept of "debt." As a result of the respondent's interpretation outlined in this summary, this issue was coded as CN because of the negative cultural value attached to the concept of "debt" in Chinese. A mortgage might not be considered a "debt" partly because debt is seen as socially undesirable.

Finally, Question 7a asks about the rooms in the respondent's home, and one of the terms included was "porch." See Figure 11:

Figure 11: Question 7 from the ACS Housing.

The interviewer asked: "Have you heard of a porch before? What do you think they mean by 'porch' here?" Then the interviewer showed a picture of a porch, and said: "This is a picture of a porch. Can you think of a better way to say this?" The interviewer wrote in a summary that the respondent had not heard of the term and did not have a clear idea. "R guessed it would be like a greenhouse where you can raise plants or enjoy teas. When R saw the picture, she thought this could be called 발코니 (phonetic translation of 'balcony') or 현관과 이어지는 발코니('balcony which is connected to a porch')." This issue was coded as SP because the respondent did not have experience with the concept of a porch, at least the typical American-style porch, and needed an explanation. (In contrast, if she had had experience with porches but didn't understand or approve of the translation, then it would not be coded SP.)

In these examples, we have seen how the coding scheme for classifying translation issues can be applied both during expert review of translations as well as cognitive interviewing with respondents not familiar with the survey translation process. It is important to point out that the situations described above are examples of issues that arose in individual cognitive interviews, and were coded in accordance with the coding scheme so that questions could be evaluated across many sets of interviews. We cannot say, at this point, whether these issues apply to a critical mass of native speakers of Chinese or Korean; but what is important for understanding a respondent's interpretation of a survey is that a problem was discovered. The analysis of multiple interviews will begin to show whether issues uncovered in individual interviews are either idiosyncratic (particular to an individual respondent) or common enough to warrant a significant modification to the translation.

Next, we will discuss some of the practical uses of the coding scheme, specifically examples of some of the solutions that could be implemented as a result.

4. Discussion

The translation review process can most effectively detect issues with Translation Errors and Linguistic Forms, while cognitive interviews conducted with monolingual respondents identify issues with Cultural Norms and Social Practices, as well as common User Errors. LF issues are quite common, and CN and SP issues are most difficult to resolve, but some potential solutions have been identified and will be discussed here.

4.1 Possible Solutions

For Linguistic Forms issues, a number of solutions have been identified through these projects. A team-based approach to translation and review and careful review procedures involving multiple parties are two of the most important aspects for catching these problems early in the process. It is important for native speakers of the target languages, as well as linguists, subject matter experts, and survey methodologists, to provide their expertise.

For Cultural Norms issues, the goal is to identify and use culturally appropriate expressions, employ culture-specific communication styles, translate the discourse structure of English into that of the target language, and incorporate politeness strategies where appropriate. Also, for CN issues, it is necessary to ask whether a certain concept exists in the target culture, and if it does, how it is expressed. These questions will help to identify early on what the cause of the translation difficulty is.

For Social Practices issues, it is necessary to revisit the source materials or source questionnaires to collect as much background and contextual information as possible. From there, translations can include explanations, examples, or notes, plus clear instructions, that are culturally appropriate and helpful to speakers of target languages. Also, we recommend flexibility when translating SP issues, and we encourage the use of descriptive phrases instead of existing terminology. For SP issues, we must ask if a certain concept exists in the target culture, and if it does not, we need to think of creative, descriptive ways to translate it. It is important to ask if the translated question might measure a concept or experience that respondents have no knowledge of, because if it does, this will lead to measurement errors. So how can the concept be translated, and how can respondents come to understand the new concept quickly and clearly?

As an example, the ACS Question 4 asks about land measurement; see Figure 12:

4 How many acres is this house or mobile home on?

☐ Less than 1 acre → *SKIP to question 6*
☐ 1 to 9.9 acres
☐ 10 or more acres

Figure 12: Question 4 from the ACS Housing.

The land measure "acre" is not used in Chinese or Korean, so "acre" is not a salient concept. Certainly the idea of land measurement exists, but not the specific unit "acre." To overcome this SP issue, the solution was to add supporting information appropriate for each target language. In Chinese, a note was included that read "One acre is about 4,000 square meters" (as meters are commonly used to describe land area in Chinese). In Korean, the supporting note read "One acre is about 1,230 pyeong" (as the measurement unit "pyeong" is unique to Korean, and commonly used).

Beyond individual issues or examples, there are also more global solutions to these translation issues. One possibility is to train translators and reviewers to identify different types of issues (LF, CN, SP) so that they can be addressed more swiftly. These findings can be shared with survey questionnaire designers and sponsors where appropriate, so that the original surveys can be constructed in a way that makes them more easily adaptable to other languages. Also, it would be useful to develop a bank of terms and concepts that are commonly used but difficult to translate and to include tested solutions so that future issues can be avoided.

5. Conclusion

Obviously, almost all translation issues are manifested in language. So how is this coding scheme different from previous survey translation-review processes, and what does it add? This coding scheme views language through the lenses of linguistic forms, cultural norms, and social practices in order to enable a better understanding of the subtle sociocultural issues that arise in translations, in addition to the challenges of translating words and sentences correctly. The application of the coding scheme gives a clearer picture of the scope of problems and enables more systematic review and revision of translations.

The next steps for this project are to refine the coding scheme to make it even simpler to use, to implement a committee approach to coding (to find out where the areas of contention are, if any), to determine inter-rater reliability, and to adjust the coding scheme as necessary based on future cognitive interview data. The anticipated applications of this coding scheme are to better evaluate the quality of translated material, measure how quality improves over time, with training, etc., and identify which questions/instructions are most problematic.

Acknowledgements

References

de la Puente, M., & Pan, Y. (2004). Census Bureau Guideline: Language Translation of Data Collection Instruments and Supporting Materials. Washington, DC: U.S. Census Bureau. U.S. Department of Commerce. April 5, 2004.

Forsyth, B. H., Kudela, M. S., Lawrence, D., Levin, K., & Willis, G. (2007). Methods for translating an English-language survey questionnaire on tobacco use into Mandarin, Cantonese, Korean and Vietnamese. *Field Methods, 19,* 264-283.

Gumperz, J. J. (1999). On interactional sociolinguistic method. In S. Sarangi, & C. Roberts (Eds.), *Talk, Work and Institutional Order: Discourse in Medical, Mediation and Management Settings*. Berlin: Mouton de Gruyter.

Harkness, J. A., Mohler, P. P., & Van de Vijver, F. J. R. (2003). Comparative research. In J. A. Harkness, F. J. R. Van de Vivjer, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (p. 3-16). Hoboken, NJ: John Wiley & Sons.

Harkness, J. A., & Schoua-Glusberg, A. (1998). Questionnaires in translation. *Cross-Cultural Survey Equivalence.ZUMA-Nachrichten Spezial, 3,* 87-128.

Pan, Y., & de la Puente, M. (2005). Census Bureau Guideline for the Translation of Data Collection Instruments and Supporting Materials: Documentation on How the Guideline Was Developed. In *Statistical Research Division's Research Report Series (Survey Methodology #2005-06).* Washington, DC: U.S. Census Bureau. http://www.census.gov/srd/www/byname.html#panyuling.

Pan, Y., Landreth, A., Hinsdale, M., Park, H., & Schoua-Glusberg, A. (2007). Methodology for Cognitive Testing of Translations in Multiple Languages. In *Joint Statistical Meeting Proceedings.* Alexandria, VA: American Statistical Association.

Willis, G. B. (2005). *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA: Sage Publications.