2.0 Method

This chapter explains the methodology used for data collection in the present study. It begins with a description of participants, and reports the process and rationale for selecting a subset of participant data for later analysis. The three instruments used in the experiment are presented, followed by details of the pilot study. The chapter concludes with a description of the experiment procedure.

2.1 Participants

One hundred and twenty-seven undergraduate students participated in an experiment conducted at a large public university in California, United States. The average age of participants was 19.2 years. Fifty-nine students were enrolled in an intermediate level Spanish grammar course which required language proficiency equivalent to four years of high school Spanish, or a 3 (out of 5) on the Advanced Placement (AP) Spanish language exam. This course was not open to heritage speakers[1]. An additional sixty-eight participants were enrolled in an advanced level Spanish content course. Enrollment in this course required at least one prerequisite course in the Spanish Department or an AP score of 5. This course was required of all Spanish majors and minors, including heritage speakers.

It was necessary to reduce the participant pool according to certain fundamental assumptions of the present study. First, native language was assumed to be important for the kinds of intuitions under investigation; therefore, participants had to report English as a native language. Second, because all target stimuli were U.S. idioms, and childhood place of residence was assumed to impact idiom familiarity and use, participants had to report childhood residence in

---

[1] McLaughlin (2001) defines a heritage language learner as "a student who learns the language of his/her home or ethnic background" (p.1).

the United States[2]. Moreover, as discussed in the previous chapter, metalinguistic awareness

may impact the kinds of learner perceptions of interest in the present study. Because research

has shown that childhood bilingualism (Bialystok, 1987; Cummins, 1979; Kellerman, 1983) and

third language acquisition (Ringbom, 1986; Gibson & Hufeisen, 2003) influence and likely

increase metalinguistic awareness, the participant pool was further reduced according to

language background, as discussed below.

The intermediate level course was not open to heritage speakers of Spanish. This meant that

any participant reporting two native languages or a non-English native language had exposure to

an L3. It made little sense to exclude data from these participants, yet retain data from

participants who had learned an L3 later in life. Therefore, at the intermediate level, data

analysis was conducted on participants who reported English as their sole native language and

Spanish as the only L2. This resulted in the elimination of data from 25 participants due to

bilingual native language (12), non-English native language (7), or L3 (6). Thirty-four

intermediate proficiency level participants were included in the final data analysis.

The advanced participant pool was also modified. As with the intermediate level, data

analysis was restricted to participants reporting English as their sole native language. The

advanced course was open to heritage speakers, which meant the elimination of data from

participants reporting Spanish/English native languages (10) and Spanish as a native language

(6). As is often the case in advanced language courses, L3 contact was prevalent.   It seems that

by the time students reach advanced language courses, most of them have begun learning a third

language. For this reason, elimination of all participants with L3 experience was not possible in

the advanced group. Early L3 exposure was possible in this group from childhood bilingualism

---

[2] The majority of participants included in the final data analysis reported childhood residence in California.
Approximately one-fifth of participants (22.5%) reported childhood residence in other states.

or a native language that was neither English nor Spanish, and these cases accounted for the exclusion of data from 15 participants. Of the thirty-seven remaining participants, 11 reported later L3 acquisition, although all claimed their L3 knowledge to be weaker than their knowledge of Spanish. Elimination of data from these participants would have made the participant pool too small for valid statistical analysis[3]. Therefore, data from participants was included in the final data analysis.

2.2 Instruments

Three instruments were created for this study: a rating task instrument, a sorting task instrument, and a language background questionnaire. The rating and sorting task instruments required the selection of idioms to be used as target stimuli. First, a list of potential target stimuli was generated, and then this list was narrowed to eliminate translation equivalents in Spanish. Finally, the most frequent and familiar idioms were selected for inclusion in the final instrument. This process is described below.

Selection of items for the three idiom categories (metaphorical image, simile, and opaque) was done through extensive review of 16 electronic idiom-dedicated corpora (See Appendix 1). In addition to screening for the semantic qualities of the three categories in question, idioms had to be at least three words long, not literally translatable into Spanish, and used in the United States[4].

None of the target items was actually acceptable in translation because if target items had been translatable, L1 transfer would have been a successful option for participants. This would have been problematic because instances of successful transfer are invisible to the researcher (do learners know that something is acceptable because of positive evidence in the L2 or are they

---

[3] Gay and Airasian (2003, p.312) specify 30 as the minimum acceptable sample size for correlational research.
[4] The websites consulted typically specified "American" idioms as the content.

using L1 knowledge?). Non-translatable target items offered an opportunity to eliminate the former possibility and focus on the role of the L1.

After reviewing the idiom corpora, the three categories were populated as follows:

Metaphorical Image: 62 idioms
Simile: 40 idioms
Opaque: 40 idioms

I translated each idiom literally, and fifteen native speakers of Mexican Spanish marked word-for-word equivalents in Spanish, which I later eliminated. Although most of the eliminated idioms received multiple votes indicating their existence in Spanish, a lone vote was sufficient for removal. Once this process was completed, the following number of idioms remained:

Metaphorical Image: 40 idioms   (22 eliminated)
Simile: 24 idioms   (16 eliminated)
Opaque: 35 idioms   (5 eliminated)

Further narrowing involved determining the most frequent and familiar idioms. Forty-four native speakers of U.S. English[5] rated the remaining idioms for frequency and familiarity on a 10-point Likert scale. Raters were relatively heterogeneous geographically, with 8 participants (18.1%) reporting childhood or current residence in California, and 36 participants (81.9%) reporting current or prior residence in over 20 states throughout the United States. The average age of raters was 25 years.

To avoid participant fatigue, the 99 idioms were divided into two randomly ordered lists of 44 and 45 items each; thus, each rater reviewed only half of the idiom list. The mean rating for each idiom was calculated and the list was re-sorted into the three categories under study (metaphorical image, simile, and opaque). For each category, the fifteen idioms rated most

---

[5] This phase of instrument development was conducted in Mexico. These participants were native speakers of English from the United States living or studying in Mexico.

frequent and familiar were selected for inclusion in the experiment (see Appendix 2 for final list).

2.2.1 The Rating Task Instrument

The instrument for the rating task consisted of the 45 non-translatable target items and 45 literally translatable distractors, for a total of 90 idioms. The translatable idioms were only controlled for direct translatability, although care was taken to include an equivalent number of translatable similes. Unlike the other target stimuli, similes have a standard form which was felt to be potentially conspicuous[6].

Each idiom was translated literally and presented in bold italics in Spanish, with its English equivalent directly below, in parenthesis. A five-point Likert scale was listed to the right of each translated idiom pair, such that the participant could circle a rating. To avoid order effects, the ninety randomized stimuli were divided into three blocks and presented in different orders to each of the participant groups. See Appendix 3 for the rating task instrument.

2.2.2 The Sorting Task Instrument

For the sorting task, the forty-five target items were listed in English. Participants were asked to sort the idioms into two groups: semantically transparent or opaque. The instructions provided a simplified explanation of the distinction:

> Please sort these idioms into two groups: transparent (T), or opaque (O). A transparent idiom is one whose meaning can be figured out from its words. For example, *play with fire* gives us an image that relates to inviting danger or trouble. An opaque idiom, on the other hand, is one whose words do not give clues to its meaning. For example, the words *butter up* appear to have no relationship to the act of flattery.

---

[6] The selection process for the distractors is described in Section 2.3.

As with the rating task, the random list of stimuli was divided into three blocks and presented in different orders to each of the participant groups. See Appendix 4 for the sorting task instrument.

2.2.3 The Language Background Questionnaire

The language background questionnaire consisted of ten questions. It collected demographic data on age, sex, academic major and childhood hometown(s). Open-ended language background questions asked about native language(s), other languages spoken, age and context of first contact with Spanish, and experience abroad. A final question asked participants to rate the similarity of Spanish and English on a ten-point scale. See Appendix 5 for the questionnaire.

2.3 Pilot Study

All instruments were piloted before use in the final experiment to measure the quantity of time needed to complete the tasks and to identify design flaws. Four volunteers participated in the pilot, which was conducted at a small, private university in Mexico. All participants were exchange students, and all were native speakers of U.S. English. Two participants were enrolled in an intermediate level Spanish course, and two in an advanced course.

Before beginning the experiment, participants were told that the study involved the translation of idioms. They were instructed to focus on the Spanish rendering of each idiom as they rated. All subjects were timed. The average time needed to complete all three instruments was 11 minutes. After completion, participant opinions were solicited and feedback noted. Specifically, participants were asked to comment about the general length of the experiment, the clarity of instructions, and the instrument format.

Piloting and consultation with advisors resulted in three changes in the instruments. The most significant change was the addition of distractors. At the time of the pilot, the rating task

involved only the forty-five (non-translatable) target items. Subsequently, forty-five literally translatable idioms were included as distractors. Idioms that had been eliminated during preliminary stages were restored, now as distractors rather than potential target items. Selection and identification of the translatable idioms required collaboration with five native speakers of Mexican Spanish.

Piloting also resulted in the modification of instructions for the sorting task. Pilot participants suggested including examples in the instructions, and these were added for the final instrument. Finally, the rating and sorting task instruments were divided into blocks and presented in three different orders. Prior to piloting, stimuli had been presented in one random order only.

2.4 Procedure

The experiment was conducted with six intact classes (three per proficiency level). With permission from instructors, the experiment was conducted in the final 15 minutes of each class. Before beginning the experiment, participants were told that the study involved idioms and second language learning. They were advised that they had 15 minutes to complete three different tasks in the sequence presented. They were told that they would see a list of English language idioms translated into Spanish. They were to focus on their opinion of the way each idiom sounded in Spanish, and rate accordingly. Participants were also told that their participation was voluntary. One student present declined participation in the experiment and left the classroom before the distribution of materials.

The instruments were presented in stapled sets, with the sorting task following the rating task to preclude participants from considering transparency or opacity when rating translatability.