Measuring Balanced Bilingual Children with Sentence-embedded Word Translation

Shin-Mei Kao

Department of Foreign Languages & Literature, National Cheng Kung University, Tainan, Taiwan Email: kaoshinmei@gmail.com

> Ferenc J. Pint ér Quan Xiang Multimedia Studio, Tainan, Taiwan Email: f.pinter@computer.org

Abstract—This paper reports two studies. The first one investigates how two 10-year-old Taiwanese-Hungarian balanced bilingual twin boys translated sentence-embedded words between their native languages over a six-month period as the physical context around them changed. The second one compares how these two bilinguals and four monolingual children defined words in their native languages. The purpose is twofold: to explore the role of an active language in translation, and to propose a new measurement for bilinguals. The reaction time (RT) and accuracy of their verbal protocols were measured and analyzed. The bilingual children's word translation reveals an active use of metalinguistic skills. The physical context did affect their RT and accuracy in general and in translating concrete/abstract concepts. The two bilinguals defined words as fast and accurate as their monolingual peers in both languages, but different definition aspects were identified from the answers of the two groups. The translation/definition task appears a comprehensive measurement for bilingual and monolingual children with any combination of languages.

Index Terms—balanced bilingual, lexical access, active language, word translation, word definition, metalinguistic competence

I. INTRODUCTION

Translating requires a thorough understanding of the concepts in the source language (SL), and demands the ability of finding appropriate equivalents in the target language (TL). For many foreign language learners, translating between their native language (L1) and the foreign language(s) (L2) is a highly demanding task till a later stage of their learning process. This is dbecause adequate translating requires an individual to transfer texts/utterances equivalently not only at the linguistic level, but at various sociolinguistic levels according to a given purpose with regard to communicative functions, styles, audiences, and other factors (Nida, 1976). For many bilingual children, translating between their two native languages is part of their everyday experience (Grosjean, 1982). Two major translation tasks performed by bilingual children are: translating between the home language and the societal one(s), and translating between different languages spoken within the family. The first task often occurs in immigrant families when the parents are not competent in using the majority language of the host country. In this case, the children receiving mainstream education often translate for their language-minority parents for daily business (Malakoff & Hakuta, 1991).

The second task is usually performed by bilingual children of a multilingual family (MF), in which two or more languages are used on daily base. The languages in the pool include the different native languages of the parents (hence, N1, N2,..., etc.), the societal language (if it is different from the two parents' Ns), and even some other foreign/second languages being learned by family members (hence, F1, F2,..., etc.) (Kao, 2008). When the parents have different Ns, more than one language is used by the core and/or extended families. Consequently, switching and translating between theses languages is common practice in an MF. Those who speak more languages, often assume the role of the translators. MF children monitor and later practice this role from their early ages.

Accessing MF children has become a fairly complex issue in the field of multilingual research and education, since there are so many possible combinations in terms of family/cultural backgrounds, amount and types of exposure to Ns, kinds and natures of Ns, physical environments, and social contexts (Baker, 2011; Cunningham-Andersson & Andersson, 1999; Grosjean, 1982). We are interested in how balanced MF bilingual children translate between the two Ns, especially when the external linguistic environment (LE) changes. In addition, we also hope to verify the so-called balanced state of a bilingual by analyzing and comparing the verbal performance made by bilingual and monolingual children of similar age. Ultimately, we hope to establish a comprehensive procedure for evaluating bilingual children of any age with any language combination.

II. LITERATURE REVIEW

A. Translation, Lexical Access and Bilingualism

Translation studies were traditionally product-oriented, emphasizing the discrepancy between the source texts and the written outcomes by professional translators. Nevertheless, translators and researchers tend to agree that translation competence is the result of a developmental process that never ends. Three developmental translating hypotheses have been proposed and they all took bilingualism into account.

The first hypothesis, initiated by Harris (1977) and Harris and Sherwood (1978), suggests that bilinguals perform translation naturally in everyday circumstances without being trained for. This hypothesis proposes that bilinguals possess an innate sense of translating, which develops continuously as they grow up. However, this hypothesis ignores the fact that not all early bilinguals develop equivalent competence in their Ns as they grow older, and thus it fails to explain how the innate sense of translation stops developing for unbalanced bilinguals (see Kaya, 2007, Lörscher, 1992, and Toury, 1995 for critiques). The second hypothesis, proposed by Toury (1995), approaches translating from a socio-cultural perspective and emphasizes the critical roles of other factors, such as the translator's motivation, personality, purpose of translation, and even responses from the interlocutors. Toury (1995) explained that during the socialization process, bilingual children develop certain strategies to fill the gaps due to their not-yet-perfect language proficiency.

The third hypothesis, proposed by Lörscher (1992), suggests that once "an individual has an even partial command of two or more languages, elementary mediations between them become possible" (p. 150). Lörscher (1992, 1996) found from the think-aloud protocols that the participants translated in two different manners: sign-oriented and sense-oriented. With a sign approach, one transfers SL segments by focusing on their lexical entries, and accesses their lexical correspondences in the TL. Sense-oriented translation, on the other hand, is a process of first separating the SL sense from its sign(s) and then representing the TL sense with a sequence of TL signs. This process requires the translator to search within appropriate situational and contextual sense ranges of the source text, and also to map sense(s) between languages. Interestingly, Lörscher (1992, 1996) found that, primarily, professional translators took the sense-oriented approach, while novice L2 learners used the sign-oriented approach toward translation. As to the bilingual children who had not received much training in translation competence as a continuum with total sign-oriented approach and total sense-oriented approach at the two ends, and the translator's needs and desires for communication determine the proportion of the two approaches to be used during the process. In other words, bilingual children live and grow with these needs and desires, rather than being born with particular abilities.

Lörscher's sense-vs. sign-translation hypothesis concurs with the two-level translation process proposed by Malakoff and Hakuta (Malakoff, 1992; Malakoff & Hakuta 1991). Malakoff and Hakuta (1991) pointed out that a translator must "apprehend and convey the meaning of the source language text, [and then] must formulate an appropriate targetlanguage sentence structure in which to embed this meaning" (p. 149). Therefore, appropriate translation requires not only one's linguistic knowledge of the two languages, but also one's ability in reformulating a message from SL to TL. Malakoff and Hakuta (1991) suggested that the successful outcome of reformulating between two languages is the result of proper application of one's metalinguistic skills. The concept of metalinguistic competence was clarified by Bialystok (1991, 2001), who defined it as the sense and ability to systematically analyze the structure of a language and to properly control the process of applying the structure in language use. This is a higher domain than oral and literate use of a language in terms of cognitive demands (Bialystok, 2001) and may develop independently in a bilingual's two Ns (Cummins, 1991). Bialystok (2007) also proposed that bilingual children need to develop a cognitive mechanism from their early age to control their attention on the production process of their Ns. Lising (2008) supported the critical role of metalinguistic ability in bilingual translation because translation utilizes all modes of reformulating a message from one language to another. Lising (2008) found that his subjects' reading comprehension on a text increased after they translated it from one language to another.

Psycholinguists have long been interested in how concepts are decoded, processed, and articulated. One influential speech production theory was proposed by Levelt and his colleagues (Levelt, 1989; Levelt et al., 1991; Levelt et al., 1999; Meyer, 1992). Levelt's theory, depicting monolingual lexical access by native speakers, incorporates a three-level feedforward activation network. The three levels from top to down are: (1) 'concept stratum', a single collection place for our conceptualized experiences; (2) 'lemma stratum', the storage place of lexical entries along with applicable inflections; and (3) 'form stratum', representing morphemes and their phonemic segments (Levelt et al. 1999, p. 4). According to this network, when a speaker intends to produce a word, she/he first gets into the conceptual phase in which relevant semantic choices are activated until a selection is made. The selected concept then enters the lemma stratum, where the syntactic features of the lexical entry, such as speech part, gender, and all possible inflections get activated. Having selected the syntactic word, the speaker encodes its morpho-phonological makeup, metrical shape, and segmental makeup before articulation sets forth. Most importantly, Levelt suggested that natural languages are processed along universal constrains. De Bot (2000) adapted Levelt's model for L2 lexical processing and suggested that there should be a separate formulator and a lexicon for each language under one large system which stores all information, with linguistic labels for all the languages. To verify speech production process across languages, Bates et al. (2003) conducted a cross-linguistic study on timed naming 520 selected picture stimuli in seven languages (including Hungarian and Mandarin Chinese), and found that lexical concepts were more or less equally accessible across the seven languages, regardless word structure and frequency.

Kroll and her colleagues proposed a "revised hierarchical model", depicting L2 learners' asymmetric translation performance from L1 to L2 (known as forward translation) versus from L2 to L1 (known as backward translation) (Kroll & Stewart, 1994; Kroll et al., 1998; Kroll & Tokowicz, 2001). This model suggests a triangular connection between L1, L2, and a common concept area. Words in L1 and in L2 are interconnected by lexical links directly, or through the media of concepts, indirectly. Kroll and Tokowicz (2001) suggested that lexical links are stronger from the backward direction, but the conceptual links are stronger from the forward direction. This model is based on comparison between the participants' reaction times (RTs) of naming pictures in both L1 and L2, and their RTs of translating single words in both directions. Many studies had been conducted in similar manner with participants of various L2 proficiency levels, language combinations, and tasks, and generally agreed that as the participants' L2 proficiency level increases, the asymmetry between the forward and backward translation diminishes (de Groot et al., 1994; de Groot & Poot, 1997; Ferr é et al., 2006). Research findings also suggest that translation is faster for concrete than for abstract words (de Groot et al., 1994) and that concrete words were easier to learn and less susceptible to forgetting than abstract words (de Groot & Keijzer, 2000). However, counter evidence was found by La Heij et al. (1996), who suggested no difference between the two directions in single word translation. La Heij et al. (1996) proposed that both translation directions were conceptually mediated; only concept activation was easier for L1 than for L2 words.

Note that the participants in these studies were mostly L2 learners with various L2 proficiency levels, learning backgrounds and combination of languages. Since the L2 proficiency was evaluated with different means across these studies, comparing these results is difficult. In addition, psychological experiments of this type usually use single words/pictures without linguistic contexts, or nonsense/pseudo words paired up with real words in the subjects' native language as the stimuli. Though these stimuli are well controllable for laboratory experiments, they do not reflect communication purposes, and restrict the range of applicable source words.

B. Language Modes and External Contexts

In terms of speech production, bilinguals, though possessing the ability of speaking two (or more) languages, have to decide which language(s) to use at a given time with a particular context. Two theories have been proposed to explain the process: Green's inhibitory control model (Green, 1998; 2000) and Grosjean's language mode continuum model (Grosjean, 2000; 2001). Green (2000) suggested that bilinguals do not switch on and off a language randomly, but that their languages have different levels of activation in the lemma system. Whether one language is selected depends on if it plays an active role in the on-going process of language production. The activation levels are controlled by the linguistic resources available. Insufficient resources may result in production errors, switching from one language to another, or mixing the language modes as "the state of activation of the bilingual's languages and language processing mechanisms at a given point in time" (p.3). Grosjean (2001) proposed that bilinguals can decide to move toward the monolingual mode by activating only one of the two languages as the base language; they can also move toward the bilingual mode with both languages activated. An example of activating the monolingual mode is when a bilingual sticks to one language when speaking to a monolingual; the bilingual can also activate both languages when speaking to another competent bilingual with the same Ns.

Green's theory is built upon the linguistic availability in one's lemma structure, while Grosjean's model explains the bilingual behavior of switching languages by the bilingual's perception toward some external contexts, such as physical locations, appropriateness for particular topics, formality, and interlocutors' backgrounds. To verify these two different views, Dewaele (2001) invited 25 adult trilinguals with Dutch as their native language and French and English as their common foreign languages to take part in two types of interview in French: casual conversation and formal oral exam. The results show the formality of the situation was a critical factor for the participants in deciding what language mode to choose. During the formal interviews, fewer code-switches were identified, suggesting a move toward the monolingual mode. This study supports Grosjean's model rather than Green's theory from the findings that language learners consciously monitor and activate certain language(s) according to external contexts.

C. Measuring Bilingualism

How to assess bilingual children and identify their proficiency levels of the claimed Ns has been a concern among educators and parents. Bilinguals are measured for a variety of purposes: placement, distribution, selection, and so forth. In many cases, bilinguals are measured with standardized tests originally designed for language learners (for example, IELTS in the UK, see Baker, 2011). The standardized tests usually measure language skills and predict literacy in one particular language, but may not be suitable for young children, especially who have not developed adequate literacy in their Ns. Neither do these tests take into account the test-takers' various cultural and linguistic backgrounds.

To overcome these disadvantages, some bilingual measurements have been designed and implemented, for example, the revised version of Woodcock Language Proficiency Battery, Bilingual Syntax Measure I and II (BSMI & II), IDEA Proficiency Tests (IPT), the Bilingual Verbal Ability Tests (BVAT), Woodcock-Muñoz Language Survey (WMLS, Woodcock & Muñoz-Sandoval, 1993) (see Gindis, 1999; Vald & Figueroa, 1994 for reviews of these tests). These standardized measurements employ procedures, such as picture naming, verbal analogies, letter/word identification, and

dictation, to evaluate test-takers' oral vocabulary, synonyms, reading, spelling, and other areas related to cognitive academic language proficiency. These standardized tests can be used with very young children; however, they also require strict administrative procedures and are not available for an arbitrary combination of languages (such as Chinese and Hungarian). Recent studies call for alternative measurements, including various kinds of surveys (Baker, 2008; 2011) and discourse analysis on speech/interaction data (Thordardottir et al., 2006; Wei & Moyer, 2008) to better portrait a holistic picture of bilinguals from the aspects of linguistic contexts, social backgrounds, self-perception, and authentic language use.

Laija-Rodr guez et al. (2007) suggested that bilingual measurements must not assess the test-taker's Ns separately, because cross-linguistic factors could influence language production in either direction and thus lead to biased interpretation of the test-taker's real proficiency in the Ns. According to Green (2000), translation is a unique measurement for bilinguals because when translating bilinguals must access the two languages simultaneously, while during the production of one language, bilinguals need to inhibit one of the Ns to avoid code-switch.

III. THE STUDIES

This paper reposts two studies. The first study investigates how two Taiwanese-Hungarian 10-year-old twin boys translated sentence-embedded words in their two Ns in different LEs. We are interested in answering the following questions. How can the translation be adequately assessed and systematically quantified? How well do they translate in each direction in terms of RT and accuracy in general. Does word concreteness play a role in their performance? What impact does the LE have on their translation performance? What do we call directions at all?

The second study investigates how two Chinese and two Hungarian monolingual children defined the same set of words in their native languages, compared to how the two bilingual children defined words. Our two goals are to set a comparison basis for assessing bilingual children and to provide justification for assessing the balanced status of MF children's proficiency in their two Ns.

A. Study One: Bilingual Translation

The participants

The two participants, Levi and Oli (their nick names), are non-identical twin brothers, born in a Taiwanese-Hungarian multilingual family. Before 4 years of age, they lived together with their parents in Taiwan with once-a-year summer visit to Hungary. Their parents adopted the "one-parent-one-language" principle within the family since they were born. Their Taiwanese mother spoke Mandarin Chinese (hence, Chinese) and their Hungarian father spoke Hungarian to them. After 4, they took turn living with their parents in Taiwan and with a host family in Hungary for the purpose of receiving formal education at regular kindergartens and primary schools in both countries. While they were in Taiwan, their active language was Chinese. There was no other Hungarian speaker nearby, so they only used Hungarian with their father. While they were in Hungary, they spoke Hungarian exclusively and maintained their Chinese with their mother via once-a-week Internet conversation. Some special arrangements were made to help the two children cope with the change of schools and living environments with cooperation from the school authorities in the two countries, and from the Hungarian host family. Their school reports show that they possessed native proficiency in both languages across all academic subjects corresponding to their age.

The contexts

The study took place during a period when the children exchanged their base countries. Three test sessions were held with an interval of two months. Session 1 was held when Levi had stayed continuously in Hungary, and Oli in Taiwan, for about one year and ten months. The external LE and the active language for Levi and Oli were Hungarian and Chinese, respectively. Session 2 was held two months after Session 1. The external LE and the active language remained the same for both children. Session 3 took place two months after Session 2. Session 3 was remarkable because it occurred three weeks after the two children exchanged their base countries. Thus, the external LE and the active language for Levi became Chinese and for Oli, Hungarian.

Each of the three test sessions contained two similar tests: a Chinese-to-Hungarian test (hence, C-H test), and a Hungarian-to-Chinese test (hence, H-C test). After conducting the tests, we realized that C-H and H-C do not represent directions themselves due to the subjects' balanced proficiency. The language used in the society provided an LE, which determined the active language of the children. When they translated from the inactive into the active language, we call it a "translating-into" task; for the other way around, we call it a "translating-away" task. For example, an H-C test in Session 1 for Levi was a "translating-away" task, but a "translating-into" task for Oli. Tables I and II summarize the relationship between the external LE and the tasks performed by the two children. The "translating-away" tasks are *italicized* in the tables.

THE RELATIONSH	IP BETWEEN THE EXTERNAL LA	ANGUAGES USED BY LEVI AND	THE TRANSLATION TASKS
Session	Session 1	Session 2	Session 3
Test type	External language Hungarian	External language Hungarian	External language Chinese
H-C Test	Translating-away	Translating-away	Translating-into
C-H Test	Translating-into	Translating-into	Translating-away

TABLE 1

THE RELATIONS	HIP BETWEEN THE EXTERNAL L	ANGUAGES USED BY OLI AND	THE TRANSLATION TASKS
Session	Session 1	Session 2	Session 3
Test type	External language Chinese	External language Chinese	External language Hungarian
H-C Test	Translating-into	Translating-into	Translating-away
C-H Test	Translating-away	Translating-away	Translating-into

TABLE 2

Instrumentation

To reflect the complexity of translation for communication, we designed a sentence-embedded word translation task, which ensured the initialization of a word within the concept stratum. This design allowed us to use words without ambiguity, caused by one-to-multiple equivalents across languages. Each test contained 18 sentences that were read to the child one at a time and the source word for translation was repeated afterwards. The child then verbally translated the word from the SL into the TL. No graphic or written representations were involved during the tests. If the children did not give any response within the initial 15 seconds, they were encouraged to translate the entire sentence. A brief follow-up interview was conducted after each test session. In the interview, the child was encouraged to explain how he reached some of his answers, and to discuss his problems with the words he was not certain about. The tests and the interviews were audio-recorded for analysis and scoring.

The 18 questions contain two even groups of words: concrete and abstract concepts. Extracts 1 and 2 exemplify an H-C and a C-H test, respectively. All the source words were selected from the textbooks used by the two schools in their previous grades to assure word comprehensibility and frequency for age-matched native children of the two languages.

Extract 1. An example question in H-C Test 3.

Question: A rendőrök k äteless ége, hogy a biztons águnkra vigy ázzanak. (k äteless ég)

Literal meaning: The <u>duty</u> of policemen is to protect our safety.

Source word: k äteless ég (duty) Abstract

Extract 2. An example question in C-H Test 2.

Question: 螢火蟲是森林裡的燈籠 (燈籠)

Ying2-huo3-chong2 shi4 sen1-lin2 li3 de0 deng1-long2. (deng1-long2)

Literal meaning: Lighting bugs are the lanterns of the forest.

Source word: 燈籠 deng1-long2 (lantern) Concrete

The Scoring System

Each question was scored in two aspects: the RT and the accuracy of the answer. Following the trend in analyzing natural speech, the RT was measured by 1/100 seconds and was represented by 1/10 seconds accuracy. In the case of no response in the first 15 seconds, the RT was counted as 15 seconds. To measure the accuracy of the answer, a fivepoint scale was developed, with 5 as the highest and 0 as the lowest score. Each answer was scored using the following five binary categories: (a) part of speech, (b) opposite, (c) self-made, (d) sentence-fit, and (e) specificity.

Category (a), part of speech, adds one point to the score if the answer's speech part corresponds to that of the source word. Category (b), opposite, adds one point to the score if the answer does not involve the antonym of the source word. In this way we penalized opposite meaning, double negation, and avoidance. Category (c), self-made, adds one point to the score if the answer does not contain words made up by the child. Making up words is often the consequence of word-by-word translation, which is frequently found among novice L2 learners (Lörscher, 1992). Category (d), sentence-fit, adds one point to the score if replacing the source word with the answer yields a native-like sentence, even if it differs from the intended answer. Category (e), specificity, adds one point to the score if the answer belongs to the correct semantic class and carries the exact semantic features of the source word. To sum up, categories (a), (d), and (e) are of "rewarding" types, while categories (b) and (c) are of "punishing" types. The two researchers scored the tests independently according to their mother tongues.

B. Study Two: Monolingual vs. Bilingual Definition

The participants and contexts

To set a basis for measuring bilingual children and to evaluate the difficulty levels of the tests, we invited two monolingual Hungarian and two monolingual Chinese children of similar age to perform a definition test (denoted by H-H and C-C, respectively). Levi and Oli also performed the same H-H and C-C definition tests. The two Hungarian children, were Levi and Oli's common friends in Hungary, while the two Chinese children were Levi and Oli's classmates in Taiwan. Thus, there were four participants for the H-H tests, and four for the C-C tests, respectively. The tests were conducted 6 months after Session 3 of Study One and the procedures were audio-recorded for scoring and analysis. The four monolingual children were evaluated as "above-average" students according to their school reports, and thus their average performance was used as the basis for comparison with that of the bilingual children.

Instrumentation and the scoring system

The questions of the three translation tests used in Study One form the definition test, except the target words were to be defined, instead of being translated. Thus, the H-H and the C-C tests contained 54 questions each, with two even portion of concrete and abstract concepts. The sentences were read to the children with the target words repeated at the end, and then the children defined the target words. Each child went through a short training session before the actual test took place. Their answers were evaluated with a similar 5-scale scoring system with the categories of part of speech, circular (instead of opposite), self-made, class (instead of sentence-fit), and specificity. The RTs were also measured in a similar manner as did for the translational tests.

IV. RESULTS

The quantitative results of Study One are presented in Sections A to D, and those of Study Two were presented in Sessions E and F. Session G reports a qualitative analysis of the two studies based on our observation during the tests and interviews with the children after the tests.

A. RT vs. LE in Translation

To understand the role of LE in the translation process, we grouped the RT results of Oli and Levi based on the "away" and "into" contexts described earlier and present the comparison in Fig. 1. Note from Fig. 1 that Oli and Levi needed longer time to perform the translating-away than the translating-into tests. The longest translating-away time (7.0 sec) and the maximal RT difference (3.0 sec) are observed in Session 2 when the children had been away from one of their LEs for the longest period of time (i.e., two years). It is also interesting to note that the smallest RT difference occurred in Session 3 (0.1 sec), which was held three weeks after they exchanged the LE. Here the active and inactive languages exchanged roles, and their RTs toward the two languages were converging. We may assume that the almost equal RTs to the two Ns in Session 3 indicate a perfectly balanced translating state for the two children.



Figure 1. Average RTs in different linguistic environments.

To verify the reason of the RT difference, we asked the two children to comment on their performances. They explained that they encountered more difficulty in searching for appropriate expressions in the inactive language than in the active one, though they knew the meaning of a particular word in SL. Oli emphasized that he forgot how to say the intended word in TL, but he knew what it meant in SL. Moreover, he could adequately translate the entire sentence into TL after he was encouraged to do so in the second attempt. He also said, when translating an entire sentence, the meaning of the forgotten word "suddenly came back" (using his words).

B. Accuracy vs. LE in Translation

In general, the two children produced quite appropriate translation in both directions, regardless the LE (see Fig. 2). They achieved an average score of 4.6 on the into tests and 4.2 on the away tests, which requires sophisticated skills and knowledge in both languages. The difference between the into and the away tests was 0.1 in Session 3, indicating an almost perfect balance. In general, their accuracy pattern is in concert with that of their RT performance. Note that the scores in Session 2 were lower than expected. Certainly, higher accuracies would show even closer relationship with the RT results, but with this small number of tests and participants, weaker characteristics may show up as opposite. We felt that unanswered questions were the main contributors of distortion; however we did not remove questions retroactively.



Figure 2. Average accuracy scores in different linguistic environments

C. Accuracy across Five Subcategories in Translation

Fig. 3 presents the results with another type of decomposition by listing the scoring subcategories, namely, part of speech, opposite, self-made, sentence fit, and specificity. The graph displays how much these linguistic areas were preferred by the two children in into/away translation. The average score for an item in the into/away tests is the sum of the average percentage values divided by 100. Similarly, the average accuracy difference is the sum of the average subcategory difference divided by 100. Since here we were interested in the long term behavior in the into and away directions, but not that when LE changed, we averaged out the scores for Sessions 1 and 2, while omitting the scores for Session 3. The two children's performed better in the first four subcategories than in the fifth one. Their accuracy rate is above 90% across the first four subcategories in the into tests, and is above 80% in the away tests. However, the accuracy difference (6.9%) in "sentence-fit" suggests that they were highly alert to sentence structures. The largest average difference in "self-made" indicates that the bilinguals may frequently turn to this strategy when translating into their inactive language.

Specifying details was the weakest area in both directions of word translation for these two children. A very great gap between the into and away context is also detected in this category. To achieve high accuracy in specificity, one must know many similar concepts within the same semantic class, and then must be able to identify subtle differences between them in both the SL and TL. It seems that such ability was still under development for our two participants.



Figure 3. Long term percentage of accuracy for the scoring subcategories in Sessions 1 and 2 together.

D. Accuracy of Concrete and Abstract Concepts vs. LE in Translation

Fig. 4 presents their accuracy scores in translating abstract vs. concrete concepts across the three sessions. They performed better in translating concrete concepts in the into tests in the first two sessions; however, they also suffered bigger declines in this aspect as they were away longer from one LE. This is indicated from a 1.0 gap in translating concrete concepts in Session 2. Interestingly, though they did worse on abstract concepts in Session 1 and 2, the differences between the into and away tests were also smaller (0.3 and 0.2, respectively) than those for the concrete concepts (0.5 and 1.0, respectively). This suggests that abstract concepts, though more difficult to learn and to express, may retain longer and more stably regardless of the LE. This finding is different from the results of de Groot & Keijzer

(2000) who suggested that concrete concepts were easy to learn and could be retained longer for L2 learners. Note that de Groot & Keijzer (2000) gave the retain test to their subject with a one month gap, while the two children in our study left one of their LEs for about 22 months in Sessions 2 and 24 months in Session 3, respectively. The learning approaches for the participants in the two studies are also different. Theirs studied nonsense words paired with L1, but ours studied concepts within meaningful contexts for communication. The two children's performance in both concrete and abstract translation became almost equal in Session 3. This indicates that LE is an influential factor for activating both abstract and concrete concepts.



Figure 4. Average accuracy scores for translating abstract vs. concrete concepts in different linguistic environments.

To further explore how translation was performed for abstract/concrete concepts, we decomposed these two areas by the five subcategories in scoring by combining their performance in Session 1 and 2. Fig. 5 shows that concrete words suffer measurably under all subcategories, especially in terms of speech parts (average difference=16.3%) and self-made (average difference=16.2%), but abstract words appear to be much more stable, especially in terms of sentence-fit (average difference=0). This is probably because abstract concepts enjoy more freedom of being substituted by descriptive expressions than concrete concepts do. The big gaps between the into and away tests across all five subcategories in translating concrete concepts suggest that the children suffered from bridging the TL and SL when one of them was not actively used for a long time.



Figure 5. Long Term percentage of accuracy for the scoring subcategories for translating abstract vs. concrete concepts in Sessions 1 and 2 together.

E. RTs and Accuracy in Defining Words between the Monolingual and Bilingual Children

The monolingual children's average RT and accuracy rate in Chinese and in Hungarian help us set the bases for evaluating the bilingual children's language proficiency. The average RT and the average accuracy scores of the monolingual and the bilingual groups are summarized in Table III and IV. In terms of language, Hungarian group performed similarly with the Chinese group (8.2 sec vs. 8.1 sec). The monolingual Hungarian children performed slightly faster (7.7 sec) than Levi and Oli (8.7 sec), while the monolingual Chinese children responded to the questions slighter slower (8.7 sec) than Levi and Oli (7.6 sec). In terms of accuracy, the monolingual Hungarian children achieved the same rate as the monolingual Chinese children did (3.6 and 3.6, respectively). Interestingly, the two bilingual

children achieved slightly higher accuracy rate than the two monolingual groups in both languages (4.1 in Hungarian, and 3.7 in Chinese). Considering the similar RTs and accuracy rates between the monolingual and bilingual groups, we confirm that Levi and Oli possessed native-level language proficiency in Chinese and in Hungarian.

C-C
.7
.6
.1
IAN AND CHINESE DEFIN
RIAN AND CHINESE DEFIN
aian and Chinese Defin C-C
LIAN AND CHINESE DEFIN C-C .6
C-C .6 .7

TABLE 3 Average RTs of the Monolingual vs. Bilingual Children for the Hungarian and Chinese Definition Tests

F. Accuracy across Five Subcategories in Definition

A١

To further analyze how the bilingual children performed differently from the monolingual children in defining words, we decomposed their accuracy scores by the five subcategories, namely, part of speech, circular, self-made, class, and, specificity. Fig. 6 shows that all the children, regardless their language background, were weaker in providing specific details of a word. They could classify concepts (i.e., giving correct class of a word), but their ability in differentiating (or describing) differences among similar concepts of the same class is yet to develop in the cognitive system. However, the bilingual group still outperformed the monolingual group by 9.3% in giving specificity. The bilingual children also performed better than the monolingual children in every other subcategory, especially in the areas of identifying speech parts (7.4%) and avoiding using circular definitions (7.9%). The bilingual children appeared to be more sensitive about the details of concepts and paid more attention to the linguistic functions of a word in sentences. When encountering difficulties in defining a concept, they relied less on the "cheap trick" of giving circular definitions, such as "tall" as "not short", than the monolingual peers did.



Figure 5. Monolingual vs. Bilingual children's Percentage of accuracy for the scoring subcategories in the definition tests

G. Summary of the Findings

The results of the definition study show that the bilingual children responded as quickly and accurately as the monolingual groups in both Ns. This is an evidence of their balanced state in both Ns. From the subcategories in accuracy, the bilingual children were more aware of the linguistic functions and semantic classification of words and paid more attention to differentiating the details of words than the monolingual groups did. The results of the translation study show that the LE influenced the bilingual children's translating speed and accuracy. When one of the Ns was inactive for a long period (i.e., up to 24 months in our study), their RT and accuracy rate decreased as they translated

away from the active language and performed more stably with translating abstract concepts than with the concrete ones. Finally, the children's balancing state between the two Ns was detected shortly after the LE changed. In other words, the after-effect of the inactive language together with the current active language helped them achieve perfect bilingualism.

V. DISCUSSION AND CONCLUSION

A. Metalinguistic Skills Enhancing Bilingualism

From their performance in defining words, the two bilinguals showed that they had built up a sophisticated classification of concepts not only at semantic, but also at syntactic level. From their word translation results, the two bilinguals showed how they accessed, generalized, bridged, and, most importantly, posed control over their speech production. The bilinguals also rationalized the problematic concepts from available linguistic cues, and found substitutions, such as synonyms, superordinates, hyponyms, and examples, to get their meanings across. These findings support Bialystok's (1991, 2001) claim that bilinguals manage communication in their Ns by applying sophisticated metalinguistic skills which had been developed and practiced since their young age. The bilingual children differentiate themselves from their monolingual peers by showing their metalinguistic competence in language production.

B. Definition vs. Translation as Measurements for Bilinguals

Realizing the close connection between coding translation and definition, we successfully transferred the translation tests into a monolingual definition test. One the one hand, the monolingual test can be used for standardizing the difficulty level of bilingual tests against the knowledge of monolinguals. On the other, the monolingual tests can verify the test taker's state of bilingualism. Compared with the existing measurements for bilingual children, our procedures do not only take the children's two Ns into account, but also access how the two Ns are conceptually bridged and verbalized from the speech protocols. The design of embedding words with sentences reflects how language is acquired and used in real world. Since no written representation is involved, the proposed procedures are suitable for measuring young children. Last but not least, the scoring process takes a reasonable amount of time and can be quantified easily for statistic analysis. Further investigation on the design of test questions, ideal question numbers, the application of different languages, and various difficulty levels may enhance bilingual research, education, and proficiency evaluation.

C. The Role of an Active Language in Translation and Language Learning

Human translation is not a symmetric process even for very professional translators. Study One depicts the possible influence of LE in terms of activating and deactivating one of the Ns for bilingual children. The balanced state was observed three weeks after our participants exchanged the LEs, by the very small differences in RT and accuracy. It is logical to assume that the balanced state will continue for a period of time until the impact of the current LE outperforms the after-effect of the previous LE. The different reactions of the bilingual children to the abstract and concrete concepts lead us to an assumption that abstract concepts may retain longer once they are learned in both language systems, but may require longer processing time. Concrete concepts, on the other hand, receive more obvious influence from the external LE. This assumption requires further verification, since opposite findings had been reported by de Groot & Keijzer (2000).

To explain the asymmetric translating behavior of L2 learners, we propose three concept areas: "exclusive L1", "exclusive L2", and "shared concepts". L2 learners who lack cultural and communication experience, have not yet developed their exclusive L2 concepts, and therefore try to mediate between L1 and L2 via shared concepts only. This may explain the asymmetry between forward and backward translation of novice L2 learners. Continuous L2 development will help to extend exclusive L2 concepts, and so to diminish the asymmetries gradually. Early bilinguals, especially the balanced ones, have been acquiring concepts within meaningful linguistic and cultural contexts. The context-rich learning approach helps bilinguals build up a larger area of shared concepts and allow easier accesses to concepts stored in the system. The language-exclusive and shared concept areas are only speculative proposals here. Analysis of extensive think-aloud protocols of translating processes will shed light on this issue.

D. Perspective Research Directions

The two studies demonstrate that definition building is an integral process for extending a monolingual's semantic range, and defining is an effective concept activator that bridges two languages. The monolingual definition test can be utilized for measuring native proficiency of monolingual children. Further research is necessary on how defining ability can find its role in L2 learning, especially when the active LE does not favor the language taught, such as in an EFL environment. The results indicate that the balancing state occurred around an LE change. The two children's schools also commented upon that their spoken language re-activations were very speedy. Further studies that intensively measure bilingual children's proficiency change during this short period may reveal the nature of the bilingual concept structure and concept retrieval according to location and culture.

ACKNOWLEDGMENT

As the parents of Levi and Oli, we are grateful for their participation, cooperation, and efforts in learning two cultures and languages not only during these studies, but in all circumstances proposed by us. We thank with our full heart for the assistance and support from the host family in Hungary, Tainan Municipal Zhi-Kai Elementary School, Taiwan, and Szent. Mớr Katolikus Által ános Iskola, Péss, Hungary in making the educational exchange and language evaluation possible for the children in the two studies.

REFERENCES

- [1] Cunningham-Andersson, U. & S. Andersson. (1999). Growing up with two languages: A practical guide. London: Routledge.
- [2] Baker, C. (2011). Foundations of bilingual education and bilingualism (5th edition). Bristol: Multilingual Matters.
- [3] Bates, E. et al. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin and Review* 10, 2: 344-380.
- [4] Bialystok, E. (1991). Metalinguistic dimensions of bilingual language proficiency. In E. Bialystok (ed.), *Language processing in bilingual children*, 113-140. Cambridge: Cambridge University Press.
- [5] Bialystok, E. (2001). Bilingualism in development: Language, literacy, and cognition. Cambridge: Cambridge University Press.
- [6] Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *The International Journal of Bilingual Education and Bilingualism* 10(3): 210-223. doi: 10.2167/beb441.0.
- [7] Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (ed.), Language processing in bilingual children, 70-89. Cambridge: Cambridge University Press.
- [8] De Bot, K. (2000). A bilingual production model: Levelt's 'speaking' model adapted. In In L. Wei (ed.), *The bilingualism reader*, 420-442. London: Routledge.
- [9] De Groot, A.M.B., L. Dannenburg, & J.G. Van Hell. (1994). Forward and backward word translation by bilinguals. *Journal of Memory and Language* 33: 600-629. doi: 10.1006/jmla.1994.1029.
- [10] De Groot, A.M.B. & R. Keijzer. (2000). What is hard to learn Is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning* 50(1): 1-56. DOI: 10.1111/0023-8333.00110.
- [11] De Groot, A.M.B. & R. Poot. (1997). Word translation at three levels of proficient second language: The ubiquitous involvement of conceptual memory. *Language Learning* 47: 215-264. DOI: 10.1111/0023-8333.71997007.
- [12] Dewaele, J.-M. (2001). Activation or inhibition? The interaction of L1, L2 and L3 on the language mode continuum. In J. Cenoz, B. Hufeisen & U. Jessner (eds.), *Cross-linguistic influence in third language acquisition: psycholinguistic perspectives*, 69-89. Bristol: Multilingual Matters.
- [13] Ferr é P., Sánchez-Casas, R., and Guasch, M. (2006) Can a horse be a donkey? Semantic and form interference effects in translation recognition in early and late proficient and non proficient Spanish-Catalan bilinguals. *Language Learning*, 56(4), 571-608. DOI: 10.1111/j.1467-9922.2006.00389.x
- [14] Gindis, B. (1999). A review of the Bilingual Verbal Abilities Tests (BVAT): A breakthrough in bilingual assessment-or is it? Communique 27: 26-27. http://www.bgcenter.com/BVATReview.htm (Accessed 28/6/2010).
- [15] Green, D.W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1, 67-81. DOI: 10.1017/S1366728998000133.
- [16] Green, D.W. (2000). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. In L. Wei (ed.), *The bilingualism reader*, 407-419. London: Routledge.
- [17] Grosjean, F. (1982). Life with two languages. Cambridge: Harvard University Press.
- [18] Grosjean, F. (2000). Processing mixed languages: Issues, findings, and models. In L. Wei (ed.), *The bilingualism reader*, 443-469. London: Routledge
- [19] Grosjean, F. (2001). The bilingual's language mode. In J.L. Nicole (ed.), *One mind, two languages: Bilingual language processing*, 1-22. MA: Blackwell.
- [20] Harris, B. (1977). The importance of natural translation. Working Paper in Bilingualism 12: 96-114.
- [21] Harris, B. & B. Sherwood. (1978). Translating as an innate skill. In D. Gerver & H.W. Sinaiko (eds.), *Language interpretation and communication*, 155-170. New York: Plenum Press.
- [22] Kaya, B. (2007). The role of bilingualism in translation activity. *Translation Journal* 11(1). http://accurapid.com/journal/39bilingual.htm (Accessed 20/08/2010).
- [23] Kao, S.-M. (2008). Language development and cultural identity evolution of children from multilingual families in Taiwan. In A. Nikolaev, and J. Niemi (eds.), *Two or more languages: Proceedings from the 9th Nordic Conference on Bilingualism*, 94-102. Joensuu, Finland: University of Joensuu.
- [24] Kroll, J.F. & E. Stewart. (1994). Category interferences in translation and picture naming: Evidence for asymmetric connection between bilingual memory representation. *Journal of Memory and Language* 33: 149-174. doi:10.1006/jmla.1994.1008.
- [25] Kroll, J.F. & N. Tokowicz. (2001). The development of conceptual representation for words in a second language. In J.L. Nicol & T. Langedoen (eds.), One mind, two languages: Bilingual language processing, 49-71. Cambridge: Blackwell.
- [26] La Heij, W.L., A. Hooglander, R. Kerling & E.V.D. Velden. (1996). Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. *Journal of Memory and Language* 35(5): 648-665. doi:10.1006/jmla.1996.0034.
- [27] Laija-Rodr guez, W., S.H. Ochoa, and R. Parke. (2007). The crosslinguistic role of cognitive academic language proficiency on reading growth in Spanish and English, *Bilingual Research Journal* 30(1): 87-106.
- [28] Levelt, W.J.M. (1989). Speaking: From intention to articulation. Cambridge, MA: MIT Press.
- [29] Levelt, W.J.M., A. Roelofs & A.S. Meyer. (1999). A theory of lexical access in speech production. *Behavioral & Brian Science* 22: 1-75. doi:10.3115/992628.992631.
- [30] Levelt, W.J.M., H. Schriefers, D. Vorberg, A.S. Meyer, T. Pechmann & J. Havinga (1991). Normal and deviant lexical processing: Reply to Dell and O'Seaghdha. *Psychological Review* 98: 615-8. doi: 10.1037/0033-295X.98.4.615.

- [31] Lising, J.L. (2008). Translation: A bilingual metalinguistic skill that facilitates better comprehension. In A. Nikolaev & J. Niemi (eds.), *Two or more languages: Proceedings from the 9th Nordic Conference on Bilingualism*, 134-146. Joensuu, Finland: University of Joensuu.
- [32] Lörscher, W. (1992). Process-oriented research into translation and implications for translation teaching. *Traduction*, *Terminologie*, *R ádaction* 5(1): 145-161. http://www.erudit.org/revue/ttr/1992/v5/n1/037110ar.pdf (Accessed 03/0912011).
- [33] Lörscher, W. (1996). A psycholinguistic analysis of translation processes. *Meta: Translator's Journal* 41(1): 26-32. http://www.erudit.org/revue/meta/1996/v41/n1/003518ar.pdf (Accessed 03/09/2011).
- [34] Malakoff, M. & K. Hakuta. (1991). Translation skill and metalinguistic awareness in bilinguals. In E. Bialystok (ed.), Language processing in bilingual children, 141-166. Cambridge: Cambridge University Press.
- [35] Malakoff, M. (1992). Translation ability: A natural bilingual and metalinguistic skill. *Cognitive Processing in Bilinguals* 83: 515-529. doi:10.1016/S0166-4115(08)61514-9
- [36] Meyer, A.S. (1992). Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives. *Cognition* 42: 181-211. doi:10.1016/0010-0277(92)90043-H.
- [37] Nida, E.A. (1976). A framework for the analysis and evaluation of theories of translation. In R.W. Brislin (ed.), *Translation: Applications and research*, 47-91. New York: Gardner Press.
- [38] Thordardottir, E, A. Rothenberg, M.-E. Rivard, & R. Naves (2006). Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? *Journal of Multilingual Communication and Disorders* 4(1): 1-21.
- [39] Toury, G. (1995). Descriptive studies and beyond. Amsterdam: John Benjamins.
- [40] Valdés, G. & R. Figueroa. (1994). Bilingualism and testing: A special case of bias. Norwood, NJ: Ablex Publishing Corporation.
- [41] Woodcock, R.W. & A.F. Muñoz-Sandoval (1993). Woodcock-Muñoz Language Survey: Comprehensive manual. Chicago, IL: Riverside Publishing Company.
- [42] Wei, L. & M.G. Moyer. (2008). The Blackwell guide to research methods in bilingualism and multilingualism. Oxford: Blackwell.

Shin-Mei Kao, born in Taiwan, obtained her Ph.D. in foreign language education from the Ohio State University, USA. Her research interests include language assessment, bilingualism, and discourse analyses.

Ferenc J. Pint ér, born in Budapest, Hungary, is a mathematician with a Ph.D. from the Ohio State University, USA. He is currently living in Taiwan, working in the field of multimedia production.