The Pennsylvania State University
The Graduate School
College of the Liberal Arts

LEXICO-SEMANTIC INTERACTION IN CHINESE-ENGLISH BILINGUALISM

A Thesis in
Psychology
by
Benjamin D. Zinszer

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2012

The thesis of Benjamin D. Zinszer was reviewed and approved* by the following:

Ping Li
Professor of Psychology, Linguistics, & Information Sciences and Technology
Thesis Advisor

Samuel T. Hunter
Assistant Professor of Psychology

Judith F. Kroll
Distinguished Professor of Psychology, Linguistics, and Women's Studies

Janet van Hell
Professor of Psychology and Linguistics

Melvin M. Mark
Professor of Psychology
Department Head

*Signatures are on file in the Graduate School.

Abstract

The hundreds or thousands of objects that we commonly encounter in daily life are often categorized under a few names that convey some similarity between them (e.g., various types of cups), a behavior known as lexical categorization. Learning to generalize these names across many familiar and novel referents is a challenge for any developing monolingual, but managing two languages adds a new layer of complexity to this task: Languages do not always agree on how objects should be sorted into these lexical categories. Previous research has shown that simultaneous bilinguals faced with these conflicting lexical category conventions may produce a unique set of categories which represent a convergence of the two languages and permit greater internal consistency across languages for the speaker. The present study examines lexical categorization in unbalanced Chinese-English bilinguals to identify changes which may occur at the interface of conflicting category information in each language. In a picture naming task, these bilinguals categorized 67 common serving dishes (e.g,. cups, plates, bowls) in each language. Categorization patterns are compared between the bilinguals and norming data from monolingual samples. By introducing new predictors describing the learner's language history and behavior (such as length of residence and code-switching frequency), we propose a statistical model to account for shifts in bilingual lexical categorization with specific interest in the effect of L2 (English) on L1 (Chinese). This model offers insight on the relationship between language history, behavior, and monolingual-likeness as well as providing a possible explanation for non-linear developmental trajectories through the mechanism of cross-language lexico-semantic interaction.

Table of Contents

List of Figures

List of Tables

Dedication

For persistence, perspective, perfection
For KNS

Chapter 1

Introduction

Non-native speech is often marked by phonological, morphological, and syntactic features that differ from those conventionally used by monolingual speakers of the language. Some of these differences stand out as highly irregular or even difficult for a monolingual to parse, a phenomenon broadly referred to as “accent.” Less obvious, but equally important for native-like language production, are nuances of meaning associated with particular lexical choices. *Semantic accent*, as described by De Groot (2012), arises from differences in the way languages divide conceptual representations between lexemes. Ameel et al. (2005) have proposed two divergent hypotheses from the observation of cross-language differences: bilinguals maintain two distinct sets of lexical-to-conceptual mappings or bilinguals the information of each language in a shared representation. Each of these hypotheses predicts different degrees and modes of interaction between languages within the bilingual mind, and it further bears important implications about the acquisition and representation of language in general.

The diversity of word-to-meaning mappings between languages has been documented across several domains of language production (e.g., lexical categorization, causal assignment, and verb aspect marking) in many languages. Since the early studies of color categorization among Native American groups (e.g., Landar et al., 1960; Ervin, 1961), researchers have noted the particular conflict that bilinguals face in managing these conflicting systems between two languages. Later studies of lexical categorization in concrete objects (Ameel et al., 2005; Pavlenko and Malt, 2011) have documented significant psycholinguistic consequences of these cross-language conflicts and explored developmental and linguistic factors which modulate the interaction between bilinguals’ representations of each language. These discoveries necessitate a re-examination of our existing models of the bilingual lexicon, specifically the relationship between lexical representations in each language and the corresponding conceptual representations. For the remainder of this discussion, this broad notion uniting lexicon and concept will be referred to as lexico-semantics, highlighting the associative relationship between these systems.

1.1 The Bilingual Lexicon

Models of bilingual lexical interaction have been sought in earnest since at least 1984, when Potter and colleagues observed that bilingual participants completed picture naming with less or equal response time to word, supporting the Concept Mediation Hypothesis of L2 access. This conclusion yielded two important implications for bilingual lexical processing. First, L1 and L2 have direct access to conceptual representations, and ostensibly even early non-immersed learners have formed these L2-concept associations. Second, L1 and L2 lexicons access a shared conceptual representation which mediates translation activities.

The Revised Hierarchical Model (RHM; Kroll and Stewart, 1994) elaborated this hypothesis with a more complete developmental account of the relationship between the L1 and L2 lexicons. The RHM posits a longitudinally variable reliance on word and concept links, the relative strengths of which modulate recruitment of conceptual and lexical resources in translation and naming tasks. Kroll and Curley (1988, as cited in Kroll and Stewart, 1994) demonstrated that sufficiently low proficiency bilinguals rely on lexical links between L1 and L2, only transitioning to concept mediated access of L2 as these conceptual links strengthen over time. Subsequent refinement of the RHM has served to highlight the importance of this longitudinal dynamicism and bilingual competition in models of lexical development. For instance, semantically mediated backward translation has been observed and seems to indicate increasing L2 proficiency in bilinguals (Kroll et al., 2010). Kroll and colleagues also noted that the competitive underpinnings of performance in the RHM allow for training effects, by which backwards translation may be semantically mediated for early learners in cases of small vocabulary and high-intensity training. As a set of associative relationships between concept and lexicon, the RHM accounts for behaviorally discrete lexical processing preferences under a model of continuous, associative lexical development, supporting competition accounts of bilingual development.

Both Potter et al., 's (1984) Concept Mediation hypothesis and the RHM imply a shared conceptual store which propagates activation to lexicons of both of a bilingual's languages. Any bilingual will recognize, however, a potential shortcoming of this assertion: namely that conceptual representations are not organized uniformly between languages. Languages differ broadly in their segmentation of conceptual space, as has been illustrated in color naming (Landar et al., 1960), causal assignment (Wolff and Ventura, 2009), object categorization (Malt et al., 1999), and verb aspect (Papafragou and Selimis, 2010). This incongruity suggests the presence of a mechanism for simultaneously storing conceptual representations across languages and maintaining distinctions in concept-to-lexicon relationships between languages[1].

[1] One issue which is commonly raised in the discussion of lexical and conceptual representations is the Whorfian Hypothesis which asserts that language shapes the most fundamental representation of concepts and thereby dictates other conceptually-driven activities such as perception and non-linguistic categorization. While this issue has been argued extensively in the domain of color categorization (e.g., Kay and Regier, 2007; Tan et al., 2008; Regier and Kay, 2009), some evidence specific to lexical categorization of concrete objects is also available. Malt et al. (1999) demonstrated that categorization differences between Spanish, Chinese, and English participants were dramatically reduced when a sorting paradigm was used instead of a naming paradigm, limiting the role of language in the task. The present study does not address the genesis of basic conceptual representations, but rather addresses the variable association of these representations to the lexicon.

Like proficiency-modulated translation effects, this difficulty has also been addressed through the proposal of an associative system. In this case of cross-language semantic incongruity, distributed representation of lexico-semantic items permits lexical units to selectively connect with salient features, as organized by the observed statistical regularities within a concept. For instance, the English word *bottle* may have strong connectivity with certain shape and material representations in the conceptual space, guided by the tendency of native English speech to prefer narrow-necked and transparent vessels for this label. However, the Chinese gloss *píng* will have significantly lower connectivity to these features, as they are less important in determining category membership among native Chinese speakers.

Monolingual computational simulations of language have depended on distributed features from the very beginning of the connectionist movement, demonstrating the power of the distributed semantic architecture in modeling learning and language production processes (Rumelhart and McClelland, 1986). Extending these principles to bilingualism, Van Hell and De Groot (1998) proposed the Distributed Feature Model to permit unique lexico-semantic representations across languages in a single conceptual store. This model emphasized connectivity between lexemes and the basic conceptual elements which compose more complex concrete and abstract representations. In cases where translation equivalents had highly congruent semantics in two languages (such as *rok* in Dutch and *skirt* in English), many conceptual nodes would activate the lexemes in both languages, facilitating semantic judgments across language. More abstract representations such as *revenge* would not activate as many of the same conceptual units as their translation equivalents (in Dutch, *wraak*). When tested against behavioral data the Distributed Feature Model accurately predicted decreased similarity ratings and increased response times for these abstract noun stimuli over concrete noun stimuli, based on the broader language-specific activation of atomic conceptual units required to represent the abstract stimuli.

This distributed and associative paradigm for modeling lexico-semantic representation has also been applied and tested in a study of semantic closeness ratings in Chinese-English bilinguals and monolinguals of each language (Dong et al., 2005). Dong and colleagues used semantic association judgments for Chinese and English words to demonstrate that English had systematically affected the Chinese lexico-semantic representations of Chinese-English bilinguals. This result adds further evidence that bilinguals maintain shared conceptual representations between languages, permitting significant partial lexico-semantic convergence between languages while some language-specific aspects of lexico-semantic relationships were preserved.

Like the Revised Heirarchical Model (Kroll et al., 2010), the distributed model of Dong et al. (2005) also accounts for longitudinal change in lexical and lexico-semantic relationships through a system of associative connections which may change in strength as language proficiency changes. This network of associative connections characterizes lexical access as a high-dimensional and temporally dynamic competitive system. Changes in associative strengths in one part of the system (such as L2 learning, or even changes to a single lexeme) affects the behavior of any other items it competes with, such as semantic associates. A considerable body of research has been dedicated towards identifying and observing these dynamic interactions between lexico-semantic systems in bilingual development, which is briefly reviewed in the following section.

1.2 Lexico-Semantic Dynamics in Bilingualism

The interaction between lexico-semantic representations in a bilingual's languages has been subject of investigation for over fifty years, since the differences in segmentation of color were documented between Zuni, Navajo, and English languages. Landar et al. (1960, p. 379) noted the importance of bilingual status in predicting changes (relative to monolingual speakers) in the lexico-semantic patterns for color categorization. Ervin (1961) later elaborated this observation with a systematic study of color naming and response time patterns in monolingual and bilingual speakers of Navajo and English. This study revealed effects not only of bilingual status, but of monolingual agreement rates, language dominance, and relative specificity of languages. Furthermore, Ervin observed that shifts in both of a bilingual's languages were driven not only by consolidation or enrichment of that language, but by ongoing competition from the unproduced language's names for the stimuli.

Decades later, these lexical categorization differences have been observed in concrete objects (Malt et al., 1999) and the effects of lexico-semantic interaction similarly documented in bilinguals' naming patterns (Ameel et al., 2005, 2009). Consequently, theoretical accounts of bilingualism have been tasked with accounting for cross-language lexico-semantic transfer. Intuitively, L1-to-L2 transfer is explained in models of second language acquisition, but more surprisingly, a growing body of literature has described the reverse relationship: L2-to-L1 transfer, wherein naming patterns in the native language are systematically altered by information learned in the second language. In highly proficient bilinguals, both of these effects may be observed simultaneously, leading to L1-L2 mutual influence.

1.2.1 L1-to-L2 Transfer

The Unified Competition Model unites the findings of several computational models and behavioral studies at all levels of language processing to form a coherent hypothesis of first and second language development based on the over-arching connectionist principles of competitive, interacting forces (MacWhinney, 2005, 2012). The UCM predicts that semantic development in a second language learner is shaped by transfer of learned patterns and associations from first to second language. In short, "whatever can transfer will" (MacWhinney, 2005). In the conceptual domain, UCM predicts that patterns of lexical categorization are initially adopted wholesale into the new language. Over time these relationships are refined based on L2 input to reflect more native-like lexico-semantic representations.

The dominance of L1 lexico-semantic representations in early L2 production is highly concordant with the L1 mediation of L2 conceptual access proposed by the Revised Heirarchical Model and the RHM's supporting research on concept-mediated translation (see Kroll and Tokowicz, 2001 for review). Under the RHM, early learners favor word-association (or lexical-to-lexical link) for L2 production, meaning that when confronted with a picture stimulus, they are most likely to name the stimulus in their L1 and subsequently seek an L2 translation equivalent. In cases where L1 and L2 lexico-semantic representations differ, this translation may result in non-native-like lexical categorization (or semantic accent).

This early pattern of L1 lexico-semantic dominance in L2 production finds support in concrete object naming for L2 learners of English (Graham and Belnap, 1986) who used Spanish category boundaries in the assignment of English words to objects such as chairs and cups of varying dimensions. In cases where L1 and L2 are incongruent in their lexico-semantic structures, L1-like patterns may precede native like naming in the L2. In lexical categorization, this effect would be most evident in cases where L1 provides a broader, more inclusive category for many objects which are divided into more specifically named categories in L2. For example, the Chinese word *píng* is typically translated as *bottle* in English, however it includes several categories of storage vessels, such as *bottle*, *jar*, and *container* (Malt et al., 1999, 2003). In early stages of L2-English learning, a native Chinese speaker might be expected to call all of these objects by the dominant translation of *píng*, that is *bottle*, instead of using the more specific names preferred by native English speakers.

1.2.2 L2-to-L1 Transfer

Although L1-to-L2 transfer is a common experience for language learners and bilinguals, the effects of L2-to-L1 transfer may be less clear. Extreme cases of immersion and dominance switching highlight the possibility for L2 lexico-semantic patterns to alter L1 production (see the L1 attrition literature for examples: Hutz, 2004; Schmitt, 2010; Stolberg and Münch, 2010), but the interactive dynamics of L2 on L1 earlier in development remains largely unexplored. Some semantic and categorization research has documented L2 on L1 effects (Dong et al., 2005; Pavlenko and Malt, 2011; Wolff and Ventura, 2009), but a coherent account of the developmental factors underlying this change to the L1 has not yet been offered.

Russian-English bilinguals living in the United States showed effects of L2-English lexical categorization in a Russian picture naming task, across multiple (early and late) ages of L2 onset (Pavlenko and Malt, 2011). These effects are measured in the adoption of English-like naming patterns (that is, the grouping of objects under the same name) during Russian production. Although the degree of English influence varied by age (see discussion regarding predictors of transfer below), even late-immersed adults showed some degree of English-likeness in their Russian naming patterns, indicating that for these bilinguals, the lexico-semantic representations of the first language remained permeable to interaction with L2 even through adulthood.

Pavlenko and Malt (2011) did not control for language dominance and limited their investigation to L2-immersed bilinguals. However, L2-to-L1 transfer may occur after a relatively limited amount of second language experience. Notably, Dong et al. (2005) identified systematic effects of English semantics on the Chinese association ratings in native Chinese speakers. In an important departure from previous expectations, these effects were observed in students of English who were immersed in a Chinese language environment, challenging the assumption that immersion is a necessary condition for the transfer of L2 information to into L1.

The existing research on L2-to-L1 transfer implicates age of onset, immersion status, and proficiency in the modulation of language interaction. While teasing these factors apart is very difficult experimentally, some computational models have begun to investigate the interaction of language systems across developmental parameters. These parameters frequently covary and confound in real language learners, impeding the focused study of any individual parameter. Computational models offer the possibility of orthogonally varying these parameters in the input to the model, and this manipulation can reveal new relationships between variables and language development.

DevLex-II (Li et al., 2007) is a computational model of lexical learning and production based on the self-organizing feature map architecture (Kohonen, 2001). Li and colleagues demonstrated that the model's interactive learning dynamics may underlie non-linear development patterns such as the vocabulary spurt and individual differences in rate of acquisition. When trained on a set of common Chinese and English vocabulary, DevLex-II revealed that the entrenchment of L1 representations over time modulated the degree of L1's lexical interaction with L2 representations, varying the degree to which language's representations could be maintained independently of one another (Zhao and Li, 2010). Similarly, Zinszer and Li (2010a) found that, in an architecturally similar model, L1 entrenchment led to non-linear patterns of L1 decay under L2 training conditions. Non-linear age of onset effects for L1 vulnerability (such as that observed by Pallier et al., 2003) were reproduced, despite the model's learning architecture and parameters being held constant. As in DevLex-II, this model demonstrated that entrenchment of the learned words, not variation in the model's learning algorithms, modulated the degree of change L2 could effect over the L1 representations.

1.2.3 Lexico-Semantic Convergence

Findings in L1-to-L2 and L2-to-L1 transfer suggest that a bilingual's lexico-semantic representations could be shaped by each language, proportionally to language dominance. This hypothesis would be consistent with the single conceptual store model of bilingualism (e.g., the Revised Hierarchical Model), but how is such a single-store system maintained between two equally dominant native languages, such as in the case of simultaneous bilingualism? Ameel et al. (2005) proposed two hypotheses for the relationship between each of a simultaneous bilingual's languages: The two pattern hypothesis predicts that the bilingual will be native monolingual-like in each language, effectively maintaining two distinct semantic systems. The one pattern hypothesis predicts that the bilingual will have a single, merged semantic system which unifies the patterns of both languages such that the bilingual is native-like in neither language but consistent between languages.

The one pattern hypothesis conforms to the Unified Competition Model's principle for maximizing transfer between language systems, but would predict relatively low native-likeness for simultaneous bilinguals in either language. In fact, Ameel et al. (2005) found that French-Dutch simultaneous bilinguals resembled a weaker form of the one pattern hypothesis. That is, lexical categories within a bilingual overlapped more than between monolinguals of either language, but the bilinguals did not completely merge representations and maintained some of the distinctions of each language. If the maintenance of distinct categorization patterns is possible within a single, interactive and competitive conceptual system, the resulting developmental trajectory for a second language learner becomes less obviously monotonic (towards L2 native-likeness) and more dynamic as L1 and L2 mutually influence the semantic representations tempered by increasingly refined discrimination in each language.

1.3 Predictors of Lexico-Semantic Change

Many demographic and developmental properties of bilinguals have been observed to covary with lexico-semantic transfer. Specifically, length of residence (Linck et al., 2009; Hutz, 2004; Schmitt, 2010), proficiency (Dong et al., 2005), age of L2 onset (Pavlenko and Malt, 2011), and code-switching (Wolff and Ventura, 2009). The role of each of these factors in predicting lexico-semantic change has not been systematically evaluated, however, nor have the individual effects of these highly inter-dependent variables (such as length of residence and proficiency) been fully disentangled. However, several studies have provided growing evidence of the general relationship between several predictors and the consequent changes to lexico-semantic representations. These data provide a general landscape of interactions that may govern the interface of L1 and L2.

1.3.1 Length of Residence

Relatively little is known about the longitudinal effects of length of residence on language transfer, principally due to the difficulty of tracking these changes over the years-long time-course of bilingual development. However, some compelling results have been found at the boundaries of very early and very advanced immersion conditions. Short-term immersion experiences (i.e., three months) can produce measurable deficits in L1 verbal fluency tasks (Linck et al., 2009). Whether this general lexical inhibition of the L1 lexicon constitutes a reorganization or shift of L1 lexico-semantic representations is arguable, particularly given the restoration of L1 performance after returning to an L1 immersion environment. Nonetheless, this inhibitory effect may be a precursor to more lasting changes as competing connections between concept and lexicon are reinforced while the previously dominant L1 competitors are temporarily suppressed. At longer lengths of residence, more specific changes in lexical categorization have been observed. For instance, Russian-English bilinguals with over twenty years of residence in New York City have been observed to produce English-like lexico-semantic patterns in their Russian production, such as semantically merging verbs *prosit'* and *sprashivat'* that share an English translation equivalent, *ask* (Schmitt, 2010). At even longer lengths of residence, entire idioms may be transferred back into the first language (Hutz, 2004), suggesting that at this stage L2 has established a degree of lexico-semantic dominance over the first language.

1.3.2 L2 Proficiency

Length of residence and L2 proficiency are often highly correlated and difficult to isolate for experimental purposes, but in at least one case lexico-semantic transfer has been compared between levels of proficiency independently of L2 immersion. Chinese-English bilinguals who study English at the undergraduate level in China showed significantly higher L2-to-L1 transfer in their Chinese semantic associations in the third year class versus the first year (Dong et al., 2005). Along similar lines, Zinszer and Li (2010b) demonstrated that after two weeks of training with a small artificial lexicon, participants showed changes in L1 picture-naming times which were sensitive to lexico-semantic patterns of the artificial L2. In each of these studies, L2-to-L1 transfer effects varied across relatively weak or moderate L2 proficiency in the absence of L2 immersion.

1.3.3 Age of Onset

Age of onset as a predictor of language development remains a controversial topic, as evidence for and against a sensitive period for language acquisition is weighed across diverse levels of achievement and other confounding age-related variables. Early inquiry into the effect of age in second language acquisition focused on syntax and suggested a discontinuity around the age of 15 years for ultimate L2 proficiency attainment in Chinese-English and Korean-English bilinguals (Johnson and Newport, 1989). Subsequent reanalysis of these data and replication of the experiment with Spanish-English bilinguals demonstrated that proficiency is, at most, sensitive to age and not centered around a criticality point, and more plausibly ultimate attainment levels vary continuously across all ages instead of only within an window of early acquisition (Birdsong and Molis, 2001). With respect to lexico-semantic development, similar age effects have been observed in lexical categorization for Russian-English bilinguals. Simultaneous or early-onset bilinguals were observed to show the greatest L2-to-L1 transfer in their naming patterns of kitchen objects while later acquirers showed less, but still significant, levels of transfer (Pavlenko and Malt, 2011). Overall, evidence suggests a gradual decrease in the vulnerability of the L1 to transfer effects from the L2.

The cause of this age-related effect in language development remains in question as sensitive period explanations have often appealed to biological causes (most notably, Lenneberg, 1967 which posed pubescent brain lateralization as the upper limit on native-like language acquisition), but more recent computational modeling efforts (Li et al., 2007; Zinszer and Li, 2010a) have demonstrated that entrenchment of L1 representations may also produce age effects which resemble a sensitive period without variation in the underlying learning system. In either case, the question of linearly varying age-effects or a discontinuous age-related sensitivity to language acquisition remains unresolved.

1.3.4 Code-Switching

Recently, code-switching behavior has been implicated as a potential mechanism of first language change. Research in age of onset effects in L2-to-L1 transfer has highlighted the differences between children who receive regular and consistent native-like L1 input and those who use it only at home or other limited settings (Pavlenko and Malt, 2011). Children exposed to a more broadly bilingual environment may experience rehearsal benefits to L1 across contexts and reduce transfer from L2. Alternately, evidence from simultaneous bilingualism (Ameel et al., 2005) suggests that even in a highly bilingual environment, convergence between languages is high, although the degree of intra-sentential code-switching in these settings has not been evaluated with respect to language transfer.

Retrieval-induced reconsolidation, a learning mechanism derived from memory research, may promote L2 transfer through code-switching. This assertion was first offered by Wolff and Ventura (2009) who linked previous studies of sequence memory with language learning. Wolff and Ventura asserted that reactivation of L1 representations, either directly through code-switching or indirectly through highly overlapping semantic representations (such as in concrete objects) temporarily increases L1 vulnerability to subsequent change by L2 input. They tested this hypothesis by looking at L2 transfer in causal judgment task in which participants were asked to choose between verbs of varying degrees of agency to describe interactions between animated humans. Wolf and Ventura posited that activation of a concept for causality would temporarily expose semantic representations in both languages to input, resulting in English-like causal judgments in the Russian-English bilinguals, even when Russian is spoken (and vice-versa for English-Russian Bilinguals). Indeed, bilinguals showed the effects of their L2 in their preferences for assigning causation in the depicted activities. These data point to the possibility that other means of inducing retrieval, such as language-switching behavior, may influence the interaction between L1 and L2 lexico-semantic systems. Based on an associative concept of learning and memory (as is assumed in the most models of bilingualism), these momentary individual behaviors change the architecture of the lexico-semantic system over time and are likely to play a role in the longitudinal development of the bilingual lexicon.

1.4 The Present Study

The present study endeavors to address cross-language lexico-semantic interaction through a cross-sectional study of lexical categorization in Chinese-English bilinguals. Previous research to this end (Malt et al., 1999, 2003; Ameel et al., 2005, 2009; Pavlenko and Malt, 2011) has established systematic differences between language pairs such as Dutch, French, Chinese, English, and Russian in the categorization of common household objects such as containers and dishes. Like the studies described above, the present study compares the naming patterns of monolinguals and bilinguals in each language to measure the degree to which these objects are named similarly between groups and identify points of divergence. The important contributions from the present study will include the evaluation of the variables discussed above, that is, length of residence, age of onset, L2 proficiency, and code-switching behavior and how they contribute to cross-language naming patterns. Specifically, the present study includes the collection of learner and cognitive data with the goal of developing a predictive model for native-likeness in lexical categorization by bilingual speakers. By accounting for these many factors simultaneously, this analysis may advance the identification of mechanisms underlying lexico-semantic interaction from other covariates and refine existing models of the bilingual lexicon.

Based on the previous research which has attempted to address most of these factors in language change individually, a set of new predictions may be offered for their inclusion in a combined model. Elman et al. (1996) have argued for increasing parsimony in the explanation of various cognitive developoment patterns through the application of linear dynamics. Broadly, linear dynamics describes systems of multiple interacting factors which are each individually represented as continuous, monotonic functions, but when computed in concert result in polytonic or apparently staged development (see Discussion for a more detailed examination of this paradigm). Situating the present study's regression model in the domain of linear dynamics, the following linear relationships are hypothesized for the regression model of L2-to-L1 transfer:

- Age of L2 onset is predicted to have a negative main effect on L2-to-L1 lexico-semantic transfer, such that later ages of onset result in a lower mean level of change in the native language. Non-linearities in age effects such as those predicted by Zinszer and Li's (2010a) model may be explained by longitudinal interactions with other variables such as L2 proficiency.

- Length of residence in the L2 environment is expected to relate positively with L2-to-L1 lexico-semantic transfer. A case study by Hutz (2004) found exponential decay of L1 semantic accuracy over time, which may suggest interaction with other factors such as overall language proficiency.

- L2 proficiency is not expected to have a significant main effect. Lexical training frequently relies on category prototypes to characterize L2 vocabulary, and participants with no immersion experience (which compose a large portion of the sample in this study) are unlikely to have sufficient exposure to the subtle differences between categories at any level of L2 proficiency. By contrast, participants immersed in an L2 environment will have greater opportunity to learn these fine grained lexico-semantic relationships, suggesting that L2 proficiency is only an effective predictor insofar as it relates to an immersion experience. Therefore, proficiency is expected to interact positively and significantly with length of residence.

- The connection between code-switching and lexico-semantic is currently undeveloped, but theoretical investigation by Wolff and Ventura (2009) based on empirical evidence of transfer suggests that greater frequency of code-switching would make L1 representations vulnerable to transfer from L2 by means of retrieval-induced reconsolidation. Code-switching frequency is predicted to be positively related to L2-to-L1 transfer.

Chapter 2

Methods

The over-arching goal of the proposed study is to quantify the competitive dynamics of first and second language lexicons in the developing bilingual. While lexical development is a broad domain which has been studied extensively in early and late L2 acquisition, considerably less research has been committed to examining the effects of L2-to-L1 transfer on the bilingual lexicon. Lexical categorization represents one facet of language development which has been documented in monolingual and bilingual speakers on both stable and attriting trajectories. Categorization behavior has been demonstrated to be subject to the influence of both L1 and L2, even in late acquirers (Pavlenko and Malt, 2011), making it a convenient and sensitive tool for measuring the effects of each language across a diverse population of bilinguals. Previous research has established the importance of bilingual status in the organization of lexical categories in adult speakers (Ameel et al., 2005). The present study takes cross-sectional measurement of lexical categorization across several language parameters using a common set of stimuli and tasks.

2.1 Participants

Two groups of students in the United States and China participated in this study. Forty-three Chinese-English bilingual undergraduate students were recruited from the Psychology department subject pool at Penn State University (State College, PA, USA). Fifty-three Chinese-English bilingual undergraduate and graduate students were recruited by word-of-mouth and online campus message board at Beijing Normal University (Beijing, Beijing, China). A comparison of the Penn State and Beijing Normal groups is provided in Table 1. Generally speaking, the students at Penn State were slightly younger than those at Beijing Normal, were first exposed to English at a younger age, and had higher self-rated proficiencies in English.

Although many participants reported some degree of training in a third language, most rated themselves at very low proficiency. Participants who self-reported a proficiency of 2.5 or greater on a 7-point scale (averaged across four ratings: reading, writing, speaking, listening) in their third language were excluded from analysis. Additionally, participants who reported speaking a third language but provided no proficiency ratings were excluded. In total, 15 participants from Beijing Normal and 9 participants from Penn State met this criterion, and their data were excluded from further analysis. Third languages included French, German, Russian, Mongolian, Japanese, Korean, Taiwanese, and Cantonese.

Penn State students ranged in age from 18 to 23 (M=19.5, SD=1.2). They were first exposed to English between ages 1 and 16 (M=8.2, SD=3.7), and self-rated their English proficiency between 2.5 and 7.0 (M=4.7, SD=1.3). The Penn State students had resided in the United States for 0 to 19 years (M=5.0, SD=5.9). Students at Beijing Normal University had ages ranging 18 to 28 (M=22.8, SD=2.0), age of earliest English exposure was 5 to 15 (M=11.4, SD=2.1). Their English proficiency varied between 1.3 and 5.5 (M=3.9, SD=1.0), as some were studying English while others majored in different subjects. None of the participants at Beijing Normal reported living in an English-speaking country.

2.2 Materials

All participants completed a language history questionnaire (Li et al., 2006, LHQ) to assess bilingual status, L2 proficiency, age of L2 acquisition, and behavioral predictors such as patterns of code-switching. The LHQ was available in both English and Chinese (simplified script) and administered according to the dominant language environment. With respect to code-switching, the LHQ allows participants to self-rate their frequency of code-switching in four contexts: Spouse & Family, Friends, Co-Workers, and Classmates. Participants rank their code-switching frequency in each context ordinally, ranging from "Rarely" to "Very Frequently." These responses are transformed into a Likert score between 1 and 5 and averaged within context group to produce the CS scores.

Participants completed an abbreviated version of the Peabody Picture Vocabulary Test (Dunn and Dunn, 2007), adapted for computer presentation. The modified PPVT was provided as an additional measure of English proficiency with specific emphasis on lexical development, the most relevant domain for the categorization task. Most participants at Beijing Normal University scored below the minimum level (Level 12) provided in the computerized PPVT, causing a floor effect and rendering the data ineffective as a measure of relative English proficiency. PPVT was excluded from the data analysis.

Early trials at Penn State revealed that many participants fail to complete the code-switching section of the LHQ or claim to never code-switch, a self-report that may (in some cases) be influenced by stigmatization of the behavior. An additional Code-Switching Questionnaire (CSQ) was added to subsequent trials to specifically probe participants' code-switching behaviors and screen for effects of negative attitudes about code-switching. The CSQ was administered as a written document in English.

An Operation-Span (O-Span) test was also used to screen participant groups for systematic differences in working memory, a cognitive factor that might confound with language proficiency or language transfer. The O-Span includes mathematical and verbal components: participants judge the accuracy of math equations and are provided a word to remember after each judgment. After several math and word combinations, participants are prompted to recall the words they have seen. The math component was administered using Arabic numerals for the math component (consistent with both Chinese and American math education) and Chinese characters for the verbal component. Participants entered their judgments using a computer keyboard and recorded their verbal responses on a paper worksheet.

Sixty-seven photographs of common household objects were used to elicit category names from participants. These objects were drawn from a stimulus set (called the *dish set*) used by Ameel and colleagues (2005) to reveal cross-language lexical categorization differences in Dutch-French bilinguals. Photographs were displayed at 480x360 pixels on a personal computer equipped for digital recording. Each photograph was accompanied by a prompt *What is this?* in the evoked language.

2.3 Procedure

Participants at Penn State University completed the tasks in two sessions separated by at least one week, with session order counter-balanced between participants. During the Chinese session, participants performed the O-Span test, then the Chinese picture-naming task, and finally the code-switching questionnaire. During the English session, participants completed the PPVT, followed by the English picture-naming task, and finally the LHQ.

Figure 2.1. The order of tasks at Beijing Normal University and Penn State University. Order of BNU tasks was constant, PSU Chinese and English sessions were counter-balanced across participants.

Participants at Beijing Normal University were only recruited for a single session, so activities were arranged to minimize cross-language interference in the naming task, as depicted in Figure 2.1. The English tests were administered first because it seems more likely that participants would overcome the primed influence of their L2 in the time it took to complete the intervening O-Span than participants being able to overcome their primed L1 responses to produce the most native-like L2 responses possible. A Chinese LHQ was completed at the beginning of each session. Next, participants completed the PPVT in English, followed by the English picture-naming task, and then the CSQ (also in English). Following the English tasks, participants completed the O-Span in Chinese, and finally they performed the Chinese picture-naming task.

In the picture-naming tasks, participants were instructed to overtly name photographs of objects depicted on the computer, in specific contrast to naming the objects' contents as illustrated by two photographic examples: a grocery bag full of vegetables (called *bag* or *dàizi*) and a trash can full of paper (called *trash can* or *lèsè tǒng*). These instructions were provided in written form using the evoked language for the naming task. Participants were also provided two practice naming trials for photographs of unrelated bottle-like stimuli, followed by the most dominant names for each stimulus (*píng* and *tǒng*) to demonstrate the desired response type. Participants were permitted to take as long as they desired to name each picture to ensure that they selected what they considered the best name for each object. Due to storage constraints, only the first ten seconds (from the onset of the stimulus) of participants' responses were digitally recorded by the computer for each stimulus.

Chapter 3

Results

3.1 Produced Naming Patterns

Participants' voice responses to the picture-naming task were transcribed and used to construct the lexical categorization patterns of each participant. Chinese responses were transcribed verbatim by Chinese-English bilingual research assistants who did not have access to the corresponding images. The English responses were transcribed verbatim by English monolingual research assistants. When these transcriptions were complete, the experimenters isolated single lexeme names for use in the categorization patterns. In English, this process entailed identifying the head noun from each description as the object's name. For example, "a small baby bottle" would be reduced to "bottle," and "a tall cup with some milk in it" would be reduced to "cup." In Chinese, each name includes an obligatory classifier that denotes its category membership, e.g., *bēi* in "bēi zi" (*cup*) or "chá bēi" (*tea cup*). In the case of self-corrections in either language, the last response was prioritized over previous responses.

Monolingual naming patterns were previous collected by Malt (2011, personal communication) using the same set of stimuli in Kaifeng, China and Bethlehem, Pennsylvania for Chinese and English, respectively. In total, 25 monolinguals were included in the Chinese norming sample, and 28 monolinguals were included in the English norming sample. Dominant name lists were derived from these samples by taking the most frequently produced name for each object by speakers of each language. The dominant name list offers the advantage of a single name for each object with which participants' responses can be compared. In most cases (English: 73%, Chinese: 97%), agreement for the dominant name was greater than 50% among the norming samples. However, this method ignores within-language variability as it occurs even among monolinguals. For that reason, the dominant name lists are used only for participant-wise comparisons, and the entire naming distributions of the norming samples are used when possible for group-level comparisons.

3.2 Individual Monolingual-Likeness Scores

Using the head noun naming patterns for each participant, pairwise comparisons were made between each participant's produced naming patterns in Chinese and the Chinese dominant names list. Comparisons take the form of an association score between zero and one, where greater values indicate higher agreement in the way objects are grouped into categories. This score is computed by making pair-wise comparisons of every object to every other object and determining whether they share a name. If objects k and l share a name in the dominant list and also share a name in the participants' list, the association score increases. If objects k and l do not share a name in the dominant list but do share a name in the participants list (or vice-versa), the agreement score decreases. This process of comparison is captured in the formula (adapted from Romney et al., 1986 as cited in Malt et al., 1999):

$A_i = \sum_{k,l>k} \frac{X_{ikl}==X_{Nkl}}{\binom{n}{2}}$

where i is the participant index, N is the dominant name list, k and l are object indexes, and X is the test (outcome 1 or 0) of whether k and l have the same name. $\binom{n}{2}$ is combinatorial notation for the number of unique pairs which may be drawn from n objects. Under ideal circumstances, n is 67 (all objects), but in many trials, no name is available for the object due to recording quality or participant error. For this reason, n is limited to the number of objects for which a participant responded. To minimize the effect of non-response items, participants who responded in fewer than 80% of trials were entirely excluded from the analysis, leaving a total of 63 Chinese-English bilingual participants.

Association scores between each participant and the Chinese norm were bimodally and approximately normally distributed around means 0.56 and 0.856. To provide some context for these scores, the association between the English and Chinese dominant names lists was also computed, with a score of 0.76. This comparison establishes the baseline level of agreement between monolinguals of each language on the way objects in the set should be categorized. Any transfer between English and Chinese lexico-semantic systems should result in an association score between 1 (monolingual-like production in the produced language) and 0.76 (monolingual-like production in the other language, or the produced language's categories are completely dependent on the other language). In light of this baseline, scores significantly below 0.76 are highly problematic, suggesting the influence of some other lexico-semantic system (for instance, a third language) on the produced naming patterns of the Chinese-English bilinguals.

The possibility of a few unusually difficult objects causing the low association scores was addressed by checking naming agreement among the bilingual samples for each object. Objects for which there was less than 0.50 agreement in both languages were judged as having produced erratic naming patterns and association scores were recomputed with these twelve objects excluded. Surprisingly, exclusion of these objects had very little effect on the association scores, which remained bimodally normally distributed. A paired-samples (within participant) t-test ($t(50)$=-2.46, p=0.017) revealed that association scores were minimally affected by the removal of these low-agreement objects: a 95% confidence interval estimated that association scores only

Comparison of High and Low Monolingual-Likeness Group

Factor	"High" Mean	"Low" Mean	Signif. *t*-test
Age (years)	20.6	22.4	0.005*
AoEE (years)	9.9	10.2	0.74
LoR (years)	3.0	0.1	0.008*
EngProf (1-7)	4.4	4.3	0.75
O-Span	48.8	48.0	0.73
CSFreq (0-5)	2.3	2.1	0.74

Table 3.1. Learner variables compared between High and Low Monolingual-Likeness Groups, AoEE (age of earliest L2 exposure, LoR (length of residence in L2 environment), EngProf (self-rated English proficiency, 1-7), O-Span (Score on O-Span working memory task), CSFreq (self-rated frequency of code-switching, 0-5)

increased by between 0.026 to 0.003. Further, the two measures are very strongly correlated (R=0.969, p<0.001) indicating that while dropping the low-agreement items may have slightly raised the mean association scores, the new scores were simply a linear transformation of the old scores without providing any new information about participants' performance.

3.3 High and Low Monolingual-Likeness Groups

If the bimodality of the group is not an artifact of outlier objects, some learner factor describing the participants could qualitatively distinguish the two groups. Participants were divided into high and low monolingual (Chinese) likeness groups around the 0.76 baseline score (i.e., above or below the baseline association score). These two groups were compared for systematic differences that might predict their classification in the high or low monolingual-like groups. Table 3.1 depicts the comparison of these groups over several learner factors. Age of the participants ($t(47)$=2.97, p=0.005) and length of residence in the U.S. ($t(26)$=-2.87, p=0.008) both showed significant differences across groups, however they do not appear to be sufficient as explanatory variables. The difference in age is small with no clear theoretical connection to native-likeness (i.e., not related to age of either language onset or immersion, simply calendar age). Length of residence also showed a small difference, with longer residence terms being associated with greater monolingual-likeness. Again, this explanation does not appear to coincide with any reasonable interpretation of language development. Both of these effects may arise from the marginal over-representation of the Penn State sample in the High group and Beijing Normal sample in the Low group. Again, this division is unintuitive, but will be explored in a bit more detail in the Discussion.

If the participants are truly best described by dividing them into two groups on the basis of their Chinese monolingual-likeness, the reliability of their responses should increase within groups as opposed to across the entire sample. To test this hypothesis, the Spearman-Brown prediction formula was used to compute split-halves reliability for the object-wise naming distribution correlations produced by each half (see Ameel et al., 2005 for details of this correlational method), with halves selected at random 500 times. This procedure was carried out within sam-

Split-Halves Reliability Scores

Sample	"High" Group	"Low" Group
Penn State	0.95	0.001
Beijing Normal	0.93	0.09

Table 3.2. Mean split-halves reliability scores for high- and low-monolingual-like groups in each location, halves randomly selected 500 times.

ple groups, allowing the possibility that the high and low groups in each location differed between each other. Table 3.2 shows the mean Spearman-Brown split-halves reliability for each Sample-by-Monolingual-likeness group. The highly monolingual-like groups were highly homogeneous, with mean reliabilities 0.95 at Penn State and 0.93 at Beijing Normal. By contrast, the low monolingual-likeness groups appeared to be highly heterogeneous, with extremely low reliability, even when divided by Sample location. Finally, to confirm that the highly monolingual-like groups at Beijing Normal and Penn State were similar in composition, the two groups were correlated at R=0.93 ($p<0.001$).

For the purpose of further analysis the high monolingual-like group is of greatest interest and most suitable for investigation of lexico-semantic interaction between Chinese and English. Most importantly, the highly monolingual-like group represents the only participants that produced a plausible set of categories for Chinese-English bilinguals of any degree of proficiency. Based on the norming sample, Chinese monolinguals show extremely high agreement (0.93 based on split-half reliability) in the naming of these objects, suggesting relatively low diversity in object names among native-speakers. If highly proficient Chinese-English bilinguals were heavily influenced by English naming norms, one would expect them to have a mean association score around 0.76, representing complete dependence on English naming patterns in their Chinese production. Instead, 0.76 appears to be the lower threshold for the highly monolingual-like group, indicating that this group contains the full range of possible dependence on English categories, and the low monolingual-like group has association scores distributed much lower than the baseline. The changes in this group do not appear to be attributable to English influence, either, as their mean correlation of Chinese produced patterns to the English norm was 0.73, at or below the baseline. Further, the low monolingual-like group does not appear to be a unified divergence from the Chinese and English norming data, as evidenced by the lack of within-group reliability. The differences from the norms are not consistent between members in the low monolingual-likeness group, challenging any assertion that they have been affected by a common, unidentified factor.

3.4 Group Correlation Analyses

With the participants of interest (high monolingual-likeness group) isolated from the others, comparison of the naming distributions for these bilinguals to the monolingual norms could be conducted. This measure differs from the association scores used to categorize participants because it accounts for variability in naming within groups, using the full distribution of names in its comparison. For this reason, such scores cannot be computed on the individual participant level, but provide a richer comparison between groups and a more accurate representation of diversity in the norming samples (Malt et al., 1999). Conceptually, this method creates a correlation matrix of similarities between objects based on their naming distributions evoked from the group. Each object is described by a vector of m dimensions, where m is total number of unique names produced by monolingual participants across all objects in a given language. Each cell contains the number of monolingual participants who produced a name m for that object. The similarity score of two objects may be found by computing a Pearson's R correlation between the objects' vectors. For n objects, a group will have $n*(n\text{-}1)$ similarity scores between objects. Between-group patterns may then be compared by correlating the lists of $n*(n\text{-}1)$ similarity scores between two groups, rendering a single correlation describing the degree to which two groups agree in the way they lexically categorize the objects. A detailed mathematical account of this process is available from Ameel and colleagues (2005) in its application to monolinguals and bilinguals of Dutch and French.

Correlations between naming patterns are depicted in Figure 3.1. Circles in the figure represent a set of naming patterns from either the monolingual or bilingual samples. In the case of monolinguals, only one circle (representing the native language) is provided per group (the upper row). For bilinguals, each language is represented in a separate circle (the lower row). Lines connecting these circles indicate the correlations (computed as described in this section) between the respective naming patterns. Crucially, these scores are different from the individual-level mean association scores described in the previous sections because the present correlations account for within-group diversity. The baseline score of 0.76 between monolingual norms in Chinese and English does not apply to this analysis and is replaced by a new baseline correlation of 0.64 based on the entire name distributions evoked from each norming sample.

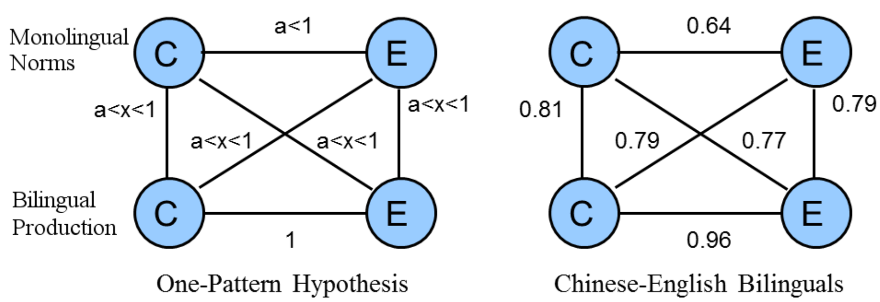

Figure 3.1. Hypothetical correlation matrix for weak one-pattern hypothesis and observed correlation matrix

According to the weak version of the one-pattern hypothesis advanced by Ameel et al. (2005), correlation between the two bilingual naming patterns should exceed the baseline correlation between monolinguals of the two languages. Further, correlations between bilingual naming patterns in one language and the monolingual naming pattern in the opposing language should exceed this baseline. Correlations were compared using Fisher's r-to-z transformation. For the Chinese-English bilinguals in this study, the convergence between naming patterns was significantly greater than between the monolinguals (Z=6.53, p<0.001). This convergence measure compares the similarity of the bilinguals' two produced languages. In this case, bilinguals' naming patterns in Chinese and English were extremely similar (0.96) compared to the baseline similarity between the two monolingual groups. This convergence score indicates only the internal consistency of the bilingual group between languages, but it does not indicate the source of these changes.

The cross-language comparisons test the respective influence of each language on the bilinguals' lexical categories by comparing the bilingual naming patterns to those of the monolingual norms. The correlation of produced Chinese naming patterns to the English monolingual norms (0.79) appeared to be greater than the baseline score (0.64). This result shows that the naming patterns produced by the bilinguals in their native language was more similar to monolingual English categories than would be expected based on the baseline similarity between the languages. However, this result was only marginally significantly greater than the baseline (Z=1.84, p=0.07). Similarly, English naming patterns to the Chinese norms higher than the baseline score (0.77) suggesting that the bilinguals' English productions bore an above-baseline similarity to Chinese categories, but this correlation fell short of significance (Z=1.46, p=0.14), failing to produce conclusive evidence of cross-language influence on the produced patterns.

3.5 Predictors of Individual Performance

To investigate the role of learner factors in the changing lexical categorization behavior of bilinguals, linear regression models were computed to measure the relative effects of the four predictors described in the introduction: age of L2 onset, length of residence in L2 environment, proficiency in L2, and code-switching frequency. Due to the great variability in task and questionnaire completion by each participant and the exclusion of the low monolingual-like group, these models had relatively few data from which to make an estimation (Chinese monolingual-likeness: 22 participants, English monolingual-likeness: 17, L2-to-L1 transfer: 22, Bilingual convergence: 10). Ultimately, none of these samples provided sufficient power to produce a significant regression estimation for their respective dependent measures.

As an alternate approach a binomial logistic regression was performed for each naming trial across participants. The binomial logistic regression estimates the probability (between 0 and 1) that any given trial will have a particular outcome. The model is estimated based on trial-wise response data, wherein the dependent measure is trial-wise response accuracy (1 or 0, compared against the dominant names list), and variables of interest are entered as predictors. This approach improves power in two ways: Individual trials are not collapsed into a single mean for each participant providing a greater number of total observations, and effects of object-wise variation in accuracy may be controlled by regressing the response accuracy over variables which describe the individual objects in each trial. By entering these object variables into the model, the effects of particular objects on the probability of answering correctly are each controlled rather than reducing the diverse object set (some of which are rarely named correctly and others which are

Binomial Logistic Model

	Chinese		English	
	Coeff.	p	Coeff.	p
Object Variables				
CH_NAMES	-0.05	0.57	-0.16	0.09
EN_NAMES	-0.41	<0.001	-0.02	0.69
CH_AGR	2.39	<0.001	-0.80	0.23
EN_AGR	0.30	0.41	3.35	<0.001
Learner Variables				
LoR	-0.54	0.004	0.79	<0.001
EngProf	0.40	0.20	0.95	0.004
AOEE	0.19	0.13	0.034	0.74
CSFreq	0.75	0.08	-3.19	<0.001
LoR:EngProf	0.14	0.007	-0.23	<0.001
LoR:AOEE	-0.02	0.03	0.003	0.75
LoR:CSFreq	0.02	0.54	0.08	0.02
EngProf:AOEE	-0.04	0.19	-0.09	0.006
EngProf:CSFreq	-0.19	0.02	0.35	<0.001
AOEE:CSFreq	-0.004	0.79	0.15	<0.001

Table 3.3. Estimated coefficients and their significance values for binomial logistic regression of trial-wise accuracy data

rarely named incorrectly) to a single association score for each participant. Each object has four variables which describe it: number of unique names in the Chinese norm, dominant name agreement (proportion of participants in the norming sample that produced the dominant name) in the Chinese norm, number of unique names in the English norm, dominant name agreement in the English norm.

Binomial logistic regressions were computed for both Chinese and English accuracy. Unlike the correlation matrices, these regressions only compare production with the dominant names of the same language, however some effect of English on the Chinese production can be measured by identifying English-specific variables which significantly predict decreased probability of accurate Chinese (monolingual-like) production. Table 3.3 lists the significance and coefficient of each object, learner, and interaction variable. In both languages, length of residence and code-switching frequency were significant predictors of naming accuracy, although their respective coefficients changed direction (positive to negative) when switching languages. Naturally, English proficiency was predictive in the English naming task. Age of onset of L2 was not significantly predictive in either Chinese or English production.

In the English production, all interactions were highly significant, except length of residence with age of L2 onset. These significant interaction terms support the general hypothesis that acquisition of a native-like L2 lexico-semantic system is a highly dynamic process which is affected not only by age predictors, but by the interactions between learner predictors over time. In Chinese production, significant interactions were fewer but depended on the variables for L2 proficiency, length of residence, and code-switching behavior. In this sense, L2 immersion and acquisition may have a detrimental effect on L1 lexico-semantic monolingual-likeness. However, the interaction between length of residence and L2 proficiency is counter-intuitively positive for Chinese production and negative for English production, suggesting that their interaction mitigates the main effects of each variable. This issue is explored in greater detail in the Discussion.

Chapter 4

Discussion

This study set out to observe the cross-language lexico-semantic dynamics of lexical categorization in Chinese-English bilinguals. Evidence of convergence and transfer between languages has been observed in simultaneous Dutch-French (Ameel et al., 2005, 2009) and sequential Russian-English (Pavlenko and Malt, 2011) bilinguals. The present study expands these findings by parameterizing likely predictors of transfer in an effort to investigate the role of learner factors in the interaction between lexico-semantic systems. The results of this study revealed strong convergence in the lexico-semantic representations used by Chinese-English bilinguals, consistent with the weak version of the one-pattern hypothesis and existing models of distributed conceptual representation for language. A logistic regression analysis further revealed significant predictors of divergence from monolingual-like naming patterns in the bilinguals. These predictors included both information about the monolingual naming distributions for objects and properties of the bilinguals which promote or protect against change in lexico-semantic representation. Further, the changes observed in the bilingual lexico-semantic system necessitate the elaboration of existing distributed models of conceptual representation, inviting further research in computational modeling.

Comparing the monolingual norming samples in each language produced a baseline correlation of 0.64 between Chinese and English. In their comparison of Dutch and French monolinguals, Ameel et al. (2005) found a baseline correlation of 0.80 for the same stimuli. This comparison indicates that Dutch and French monolinguals share a greater similarity in their lexical categorization patterns than do Chinese and English speakers. The Dutch-French bilinguals were further part of a highly bilingual community, while the Chinese-English bilinguals in this study are part of two relatively monolingual language communities: Chinese speakers in Beijing, China and English speakers in central Pennsylvania, USA. Elements of bilingual convergence and transfer observed by Ameel and colleagues which result from bilingual immersion should not be reproduced in the Chinese-English population. However convergence and transfer are observed in both populations, pointing to cognitive interaction between lexico-semantic systems, as modulated by the predictors in this study, as an underlying force in bilingual lexical categorization.

4.1 Convergence and Transfer

Group-level correlation of the naming patterns produced by the Chinese-English bilinguals in this study are consistent with the predictions of the weaker version of the one-pattern hypothesis advanced by Ameel et al. (2005). The group level correlations between languages (showing L1-to-L2 transfer and L2-to-L1 transfer) were greater than the baseline correlation between monolingual norms, although not significantly greater. More importantly, the bilinguals' naming patterns in each language showed strong convergence, with a significantly greater correlation within bilinguals (0.96) than between the monolinguals (0.64). The discovery of significant and very strong convergence within the bilinguals underscores the degree to which they are relying on a shared system of lexico-semantic representations in the production of both their native language (Chinese) and their second language (English).

The Distributed Feature Model (Van Hell and De Groot, 1998) predicted that concept-mediated translation of concrete objects would be faster than abstract words based on their greater overlap in conceptual space, a hypothesis which found empirical support. Under this model, lexico-semantic differences are driven by variations in connectivity between lexical and conceptual representations. Although the Distributed Feature Model allows binary variation in connectivity between lexemes and features, bilingual lexical categorization patterns suggest that change in bilinguals' language systems are more nuanced (Ameel et al., 2009). Conceptually represented features, such as shape, material, and color of an object are accessed by both languages, however they these languages differentially weight such features in the assignment of lexical categories. Consequently, convergence may actually be facilitated in concrete objects by the greater number of shared connections which are subjected to the respective influence of each language. In this experiment, a great deal of convergence within bilinguals' lexical categories was observed between two relatively distant languages, greater or equal to the convergence observed between Dutch and French, despite the latter pair being much closer in their monolingual categories.

The presence of the convergence effect in a unique pair of languages and between particularly distant lexical categorization systems lends support to the weak one-pattern hypothesis and distributed connectivity models of lexico-semantic representation. The similarity of the convergence effects in Chinese-English bilinguals (0.96) and Dutch-French bilinguals (0.91) may also indicate a limit to the lexico-semantic system's tolerance for inter-language variation on this set of stimuli. If the bilingual language system only supports a limited degree of lexico-semantic diversity, some minimum degree of convergence should be observed for a given stimulus set across all levels of proficiency for any pair of languages with any amount of correlation between monolingual norms. Some evidence to support this claim is found in a second set of stimuli used by Ameel and colleagues (2005) with the Dutch-French bilinguals. In their *bottles* set of 73 containers, they found a much lower degree of monolingual correlation (0.63) than for the *dishes*. Nonetheless, bilingual convergence was observed at 0.88, which is not significantly different from the convergence (0.91) observed in the much more similarly categorized *dishes* set (Z=-0.88, p=.38). A more convincing test of this hypothesis would use both the bottles and dishes sets in an extremely distant pair of languages (<0.60 monolingual agreement). For these hypothetical languages with extremely low monolingual agreement, convergence scores should still be measured around 0.90 or greater for both sets of stimuli.

Despite the robust support for bilingual convergence found in this study, the effects of cross-language lexico-semantic transfer were not sufficient to achieve statistical significance. The two-by-two correlation scores between bilingual productions in each language and the monolingual norms in each language ranged from 0.77 (the lowest, L1-to-L2 transfer) to 0.81 (the highest, L1 monolingual-likeness). Of these four, only Chinese monolingual likeness was significantly different (Z=-2.06 p=0.04) from the baseline correlation between monolinguals. The convergence score described above provides no information on the direction of the change or which language has exerted more influence on the shared pattern, but the non-significant results of the group correlations suggest that, indeed, there is transfer between languages. Proving this in the group-wise comparison would require a larger set of objects to be named or a greater degree of transfer from one language to the other. Since the convergence between both languages within bilinguals is already approaching unity, the sensitivity of the measure could be improved by focusing analysis on items which are not consistent between languages (i.e., lowering the monolingual-monolingual correlation and increasing the chance for non-monolingual-like transfer). In the individual analysis, variation in objects' categorization by monolingual speakers in the norms is accounted for in the regression, improving model's sensitivity to transfer from L2 (English) to L1 (Chinese) production.

4.2 Object-Wise Predictors

Even in balanced, simultaneous bilinguals, convergence in a limited set of concrete objects is not complete (Ameel et al., 2005), suggesting that some objects undergo a change in lexical categorization while the monolingual-like categories of others are preserved. Ameel et al. (2009) determined that while categories underwent a small shift in their conceptual centers, objects at the boundaries of categories which differed across languages were most vulnerable to change by simplification, wherein objects tended to attract toward categories of similar objects at the cost of monolingual-like naming exceptions. In the individual analyses, these differences were captured in the regression model to adjust for this effect. Four independent object-wise predictors were entered to account for variance in each objects' probability of evoking a monolingual-like response: dominant name agreement and number of total name candidates measured in each of the monolingual samples.

Dominant name agreement was a significant predictor in each language for accuracy of naming in the same language. The relationship between agreement and accuracy is self-evident, as randomly selected participants with perfectly monolingual-like patterns would produce the dominant name with the same probability as the dominant name agreement. However, under this condition, the coefficient of this predictor should be approximately one, as an increase of dominant name agreement represents an equal increase in the probability that a participant will produce the dominant name in a given trial. In the model estimated from the present data, the coefficients for same-language dominant name agreement are significantly greater than one. These larger coefficients indicate that the bilingual participants in this study amplify the influence of dominant name agreement in their own selection of a name, such that every percent increase in monolingual agreement results in greater than one percent increase in the probability that a given participant will select the dominant name. This observation is consistent with the preference for categories of like-objects, as bilingual participants are highly sensitive to dominant naming agreement and converge much more quickly (i.e., over-represent dominant name agreement) than monolinguals on a category preference for objects at the boundaries.

The number of names for an object in the opposite language was also a significant predictor ($p<0.001$) in Chinese production and marginally significant predictor ($p=0.09$) in English prediction. This effect points to the role of cross-language competition, even in dominant language production. The greater the number of candidate names (as generated by monolingual norming sample) in the other language, the lower probability that the dominant name in the produced language was chosen by the bilingual participant. Gathercole and Moawad (2010) found that early Arabic-English bilinguals were less likely to name exemplars from Arabic categories which were broader than in English than when there were one-to-one translations. Although the number of English names for a particular object in the present study derives from disagreement among the monolingual norming participants, these objects exemplify the boundaries between categories in English which are not typically distinguished in Chinese. These objects are most likely to elicit the one-to-many translation effects between Chinese and English as compared to objects which have a clearly defined category in each language because, even among monolinguals, they elicit activation of multiple English competitors.

4.3 Learner Predictors

After the effects for the differences between objects had been entered into the model, remaining variance for the accuracy of responses was regressed over the learner variables: length of residence (LoR), age of earliest exposure to L2 (AoEE), self-rated proficiency in L2 (EngProf), and self-reported frequency of code-switching (CSFreq). Each possible pairwise interaction term between these variables was also entered into the model. These learner variables were used as predictors for monolingual-like dominant name accuracy in both Chinese and English production. Because all of the learner predictors described properties of L2 learning and use in the participants, the regression of Chinese monolingual-likeness over these variables was effectively a measure of the L2's effect on native language. The regression of English monolingual-likeness over these variables was necessary to model the object-wise predictors. The learner variables in the English model also provide an account of second-language acquisition under varying conditions, which is not the inquiry of this study and will be left for further exploration at a later time.

In the Chinese model, main effects were found for LoR and (marginally significant) for CSFreq. Generally, the effect of LoR could be interpreted to suggest that as time spent living in an L2 environment increases, the probability of producing a monolingual-like name for an object decreases slightly. Conversely, as language switching becomes more frequent, there may be an effect of increasing the probability of monolingual-like production. These results should be interpreted with great caution by consideration of the significant interactions of the two variables and their interactions with other variables.

Three significant interactions were found in the Chinese model: LoR with EngProf, LoR with AoEE, and CSFreq with EngProf. LoR and AoEE interacted with a negative effect, indicating that for later ages of L2 onset, length of residence became more influential in predicting change to native language naming patterns. This finding diverges from Zinszer and Li, 's (2010a) computational model of first language attrition, wherein L1 decay rate (L1 proficiency measured as a function of LoR) was variable in early ages of L2 onset but constant in later ages of onset. In this model, however, L1 input completely ceased after L2 onset, limiting cross-language influence to effects of entrenchment rather than requiring the model to mediate conflicting inputs. For human participants who continue using both languages in varying degrees and at varying levels of immersion, this competition could result in a different interaction. Further, the main effect of LoR remains significant (and of far greater magnitude) in the presence of the LoR by AoEE interaction. In other words, length of residence in L2 environment is strongly predictive of reduction in L1 monolingual-likeness for all ages of onset, but experiences a marginal increase in importance for later ages on onset.

Frequency of code-switching interacted with L2 proficiency (CSFreq*EngProf) to reduce probability of L1 monolingual-like naming. In the main effects, neither of these predictors was significant (CSFreq only marginally so), suggesting that their importance may be limited to the higher ends of both variables' distributions. Behaviorally this interaction means that highly proficient Chinese-English bilinguals who are code-switching frequently are more likely to produce non-monolingual-like naming patterns in Chinese than those who are less proficient in English or those who do not code-switch. This interaction is opposite in valence to the marginally significant main effect for code-switching, which has a coefficient approximately four times greater in magnitude. In effect, because the interaction term is also a multiple of English proficiency, at EngProf values of 4 or greater, frequent code-switching represents a disadvantage for monolingual-like production in Chinese, while at values of EngProf significantly less than 4, frequent code-switching is an advantage.

If code-switching acts as a prime for change in linguistic memory, posited by Wolf & Ventura (2008) as a form of retrieval-induced reconsolidation, its effect on L1 will only be negative if code-switches related to lexico-semantic conflict are monolingual-like in L2 rather than L1. Due to the subtlety of these differences and the tendency of learners to rely on L1 patterns in early stages of L2 learning (MacWhinney, 2005), this code-switching can only reinforce L2-like patterns for higher levels of L2 proficiency. In cases of low L2 proficiency, code-switching would, in fact, only further reinforce L1 patterns (rather than L2 patterns) and may signal uncertainty about the name for an object in L2.

Even though a detrimental effect of LoR on Chinese monolingual-likeness was found, the interaction between LoR and EngProf is positive, suggesting that high proficiency in English mitigates the effects of LoR on Chinese production. Like the interaction between English proficiency and code-switching, the main effect of LoR is eliminated when English proficiency reaches approximately 4, suggesting that after a certain level of English proficiency, LoR ceases to be a meaningful predictor of L2-to-L1 transfer.

Non-monotonic development patterns such as those observed throughout language acquisition are often attributed to staged maturation of qualitatively distinct systems. However, these developmental trajectories can be explained as the interaction of multiple, continuous linear predictors. Using the model estimated for Chinese monolingual-likeness in lexical categorization, a U-shaped development curve might be hypothesized based on the interaction of the linear terms over time. Figure 4.1 illustrates one such scenario in which the monolingual-likeness of a Chinese-English bilingual's Chinese naming patterns are predicted over the course of a typical period of undergraduate study. While length of residence naturally increases over time, English proficiency is also allowed to increase linearly over time from 3 to 5 on a 7-point scale. Code-switching is assumed to be moderate (3 on a 5-point scale) and a typical age of earliest exposure (12 years) was selected. As depicted in Figure 4.1, the marginal probability (after accounting for object variables) of producing the dominant monolingual name for an object decreases temporarily, reaches a minimum, and increases again. In further predictions (not depicted, but computable by the reader), differing degrees of code-switching can change the shape of the function over time entirely.

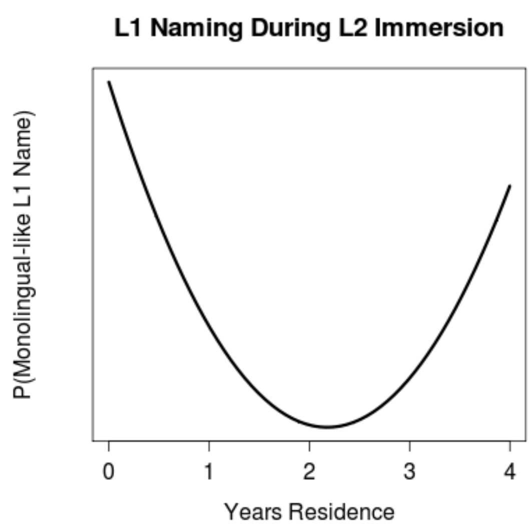

Figure 4.1. Hypothetical monolingual-likeness curve over time

4.4 Trajectories of Development

Elman et al. (1996) demonstrate that linearly dynamic systems, based on multiple interacting linear forces, can produce non-monotonic patterns, a premise which has driven the broader theoretical perspective known as Dynamic Systems Theory (Gogate et al., 2001; De Bot et al., 2007). Dynamic Systems Theory describes the sensitivity of a developing system to many inputs and the tendency of development to settle into local minima or attractor states. When several interacting predictors in a system reach a stable equilibrium or resistance to change, the system is said to have achieved an attractor state. These states are defined by their persistence until a critical degree of change has occurred in one or more parameters, breaking the system out of equilibrium. Attractor states may be manifested as developmental landmarks or stages, but they are highly variable between individuals based on differences in internal and environmental conditions (De Bot et al., 2007). This paradigm of development is highly compatible with the highly interactive models of Chinese and English lexical categorization found in this study and may extend to other areas of bilingual development that have typically been treated as discrete or independent phenomena.

The gradual and probabilistic changes in lexico-semantic representation pose a new challenge to distributed conceptual models. The Distributed Feature Model (Van Hell and De Groot, 1998) relies on the binary status of elementary features to semantically define lexemes, however as observed in concrete object naming, the presence or absence of certain features is insufficient to define category membership. Instead, features receive variable weight across languages in their importance for defining a category (Pavlenko and Malt, 2011). Similarly, these relationships are also variable over time, with defining features gaining or losing importance according to bilingual status and conflict between languages (Ameel et al., 2009). Distributed conceptual models must therefore integrate a mechanism for continuous and gradual change within existing features as well as the integration of new features. Figure 4.2 depicts one cross-language relationship in which the principal defining features (solid lines showing features common among approximately 100% of exemplars) of the translation equivalents *chair* and *yǐzi* or *sofa* and *shāfā* are highly similar (Malt et al., 2003). However, each language varies in its tolerance or occasional inclusion of other features (dotted lines showing features shared by approximately 50% of exemplars). For example, in English a seat with large cushions may be a *chair*, but the Chinese word *yizi* would no longer apply to this object. In an associative system, the dual activation of each lexical representation means that over time *yǐzi* would become increasingly associated with padded seats (see Hebb, 1949 regarding mechanism). This convergent association effect is consistent with the retrieval-induced reconsolidation hypothesis offered by Wolff and Ventura (2009). If distributed conceptual models are to account for developmental change and partial convergence, such as that observed in this study and others (Ameel et al., 2005; Dong et al., 2005; Pavlenko and Malt, 2011), they must be enriched by the inclusion of varying connection weights for partial association across diverse exemplars and over time.

Connectionist models of language learning (e.g., Li et al., 2007; Zhao and Li, 2010; Zinszer

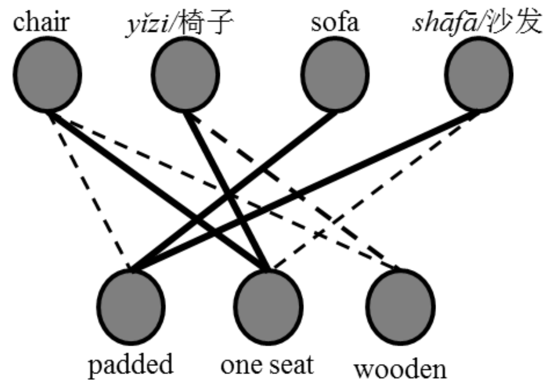

Figure 4.2. A distributed concept representation model which allows for varying connection weights. Solid lines approximate 100% association, dashed lines approximate 50%.

and Li, 2010a demonstrate that phenomena such as U-shaped development and sensitive periods can emerge from the interaction between learned information (such as morphological rules or new words). Meara's (2004) Boolean model of lexical attrition used a simple connectionist paradigm to simulate the effect of intra-lexical relationships and exhibited self-organized criticality, the wide-spread and sudden deactivation of lexical nodes at unpredictable intervals. Meara's findings may be interpreted as largely a product of the Boolean models architecture, but more importantly, Meara demonstrated the wide individual variablity in L1 change arising from the model's initial random state.

Self-organizing feature maps (SOM) have offered a promising option in modeling lexico-semantic interaction, producing several cognitively plausible models of language development (see Li, 2009 for a review; see also Richardson and Thomas, 2008; Mayor and Plunkett, 2010). The potential for extending SOM to highly parameterized problems in lexico-semantic interaction is suggested by its flexibility in simulating the effects of competing input sets. Age-related dynamic cross-linguistic competition in L2 learning has been demonstrated with other SOM-based models (Li and Farkas, 2002; Li et al., 2007). Furthermore, effects of sensitive period or catastrophic interference have also been shown with the manipulation of learning parameters in SOM (Richardson and Thomas, 2008). Computational modeling offers the possibility of a unified account of language learning and interaction, integrating empirical research under both descriptive statistical models, as proposed in the present study, and in simulating computational mechanisms for explanatory power.

One area of bilingualism which has been difficult to address experimentally and is ripe for both empirical and computational exploration is first language attrition. This phenomenon, discussed below, represents an important frontier in language development and offers a unique opportunity to test existing models and hypotheses. In the last decade, there has been an increasing amount of work devoted to the study of native language loss in individuals, L1 language attrition. Longitudinal case studies of lexico-semantic attrition have implicated length of residence in L2 environment as a predictor (Hutz, 2004; Stolberg and Münch, 2010) of these changes, which was also found in the present study. Similarly, predictors of phonological attrition have received some systematic treatment, such as age-related effects (Hyltenstam et al., 2009; Pallier et al., 2003) and code-switching behavior (De Leeuw et al., 2010). One could reasonably expect the long-term course of lexical attrition to differ from that of phonology based on differing memory stores hypothesized to support each function (Hernandez and Li, 2007; Ullman, 2001). However, the present study reveals that age and code-switching are indeed significant predictors of change in L1 lexical categorization or semantic accent in general. However, lexico-semantic attrition has been treated a primarily descriptive enterprise due to the lack of rigorous experimental methodologies for tracking these changes longitudinally. The present study and others (such as Dong et al., 2005 demonstrate that a relatively short period of time is required to observe measurable and systematic change in L1, with or without immersion conditions. The potential continuity between early L1-L2 interaction, L2-to-L1 transfer, and advanced L1 attrition should guide attrition research to more specifically target predictors that are known to affect L1 production throughout bilingual development and highlight elements of L1 production (such as lexical categorization) which are likely to be longitudinally dynamic and have known relationships with these predictors.

4.5 Conclusions & Future Research

In the present study, bilingual convergence and L2-to-L1 transfer effects have been identified in Chinese-English bilinguals across a variety of learner variables relating to their language development and behavior. These results support a weak version of the one-pattern hypothesis of bilingual lexico-semantic representation, agree with distributed models of conceptual representation, and highlight the interactive lexico-semantic dynamics in bilingual lexical categorization, even outside highly bilingual contexts such as those previously studied. Further work is needed to investigate each of the object and learner predictors implicated by the models estimated in this study. For example, the possibility that the number of L2 competitors for an object changes naming behavior in the L1 has important implications for parallel activation studies of the bilingual lexicon. The absence of main effects for age of second language onset, but its importance in interactions conflicts with common concepts of sensitive period, but may provide an explanation for apparent non-linear patterns in the relationship between age of onset and native-like language production in other domains. Further, code-switching has only recently been raised as a possible mechanism of change in bilingual lexico-semantics, and its significance as a predictor of non-monolingual-likeness in this study affirms the importance of further elaborating this relationship. Finally, the application of this model or a related model in testing longitudinal development of bilinguals has yet to be tested, but through the principles of linear dynamics, the linear predictors currently of interest may interact to explain non-linear patterns of language development. This study paints a rough landscape for the dynamic interface between L1 and L2 lexico-semantic representations, demonstrating that convergence and transfer between languages are real, measurable, and generalizable across many levels and contexts of language use.

Appendix A

Human Subjects Approval for USA

PENNSTATE

Vice President for Research
Office for Research Protections

The Pennsylvania State University
The 330 Building, Suite 205
University Park, PA 16802-3301

(814) 865-1775
Fax: (814) 863-8699
www.research.psu.edu/orp/

Date: February 11, 2011

From: Dolores W. Maney, Compliance Coordinator

To: Ping Li

Subject: Research Proposal - Modification **(IRB #34503)**
Approval Expiration Date: August 19, 2011
(Note: This date reflects the anniversary date of the actual submission approval date.)

"Cross-Language Lexical Interaction"

The revision(s) to the above-referenced study has been reviewed and approved by the Institutional Review Board (IRB). You may proceed with your study. Please continue to notify the IRB of any further changes to your study.

COMMENT: Approval of the January 26, 2011 request has been granted for the following modifications: 1. Research Personnel: Adding J. Ying from PSU for assistance with recruitment, consent, data collection, and data analysis. Those from NonPSU institutions will obtain approval for their roles in the research from their own IRB offices. 2. Inclusion: Adding Bilingual speakers of English and Chinese. 3. Exclusion: Participants who have taken more than three years (high school) or three semesters (college) of foreign language courses will be excluded from this study. 4. Funding Source: Adding NSF funding Source. Title: Now reads as Cross-Language Lexical Interaction. 5. Research Procedures: Adding a photographic picture naming task to be carried out in English and in Chinese for bilingual speakers. Adding written responses as well as verbal responses. 6. Participants: Adding 100 who are eligible to enroll. 7. Risk: No changes were noted. 8. Informed Consent Form: No changes were noted.

On behalf of the IRB and the University, thank you for your efforts to conduct research in compliance with the federal regulations that have been established for the protection of human participants.

Please Note: The ORP encourages you to subscribe to the ORP listserv for protocol and research-related information. Send a blank email to: L-ORP-Research-L-subscribe-request@lists.psu.edu

DWM/dwm
cc: Jing Yang
Benjamin D. Zinszer

Appendix B

Human Subjects Approval for China

PENNSTATE

Vice President for Research
Office for Research Protections

The Pennsylvania State University
The 330 Building, Suite 205
University Park, PA 16802-3301

(814) 865-1775
Fax: (814) 863-8699
www.research.psu.edu/orp/

Date: June 28, 2011

From: Jodi L. Mathieu, Research Compliance Coordinator

To: Ping Li

Subject: Research Proposal - Modification **(IRB #36830)**
Approval Expiration Date: May 3, 2012
(Note: This date reflects the anniversary date of the actual submission approval date.)

"Lexical representations of noun and verb in the monolingual and bilinguals"

The revision(s) to the above-referenced study has been reviewed and approved by the Institutional Review Board (IRB). You may proceed with your study. Please continue to notify the IRB of any further changes to your study.

COMMENT: Approval of the modification submitted on June 13, 2011 has been granted for the following: (1) changes to the inclusion/exclusion criteria; (2) addition of a research personnel; (3) revisions to the study procedures and data collection instruments; (4) addition of recruitment materials; (5) changes to compensation; and (6) revisions to the informed consent forms [Doc. #1001 & Doc. #1002].

Attached is/are the revised and dated, IRB-approved informed consent(s) to be used when enrolling participants for this research. Participants must receive a **copy** of the approved informed consent form to keep for their records.

On behalf of the IRB and the University, thank you for your efforts to conduct research in compliance with the federal regulations that have been established for the protection of human participants.

Please Note: The ORP encourages you to subscribe to the ORP listserv for protocol and research-related information. Send a blank email to: L-ORP-Research-L-subscribe-request@lists.psu.edu

JLM/jlm
Attachment
cc: Jing Yang
Benjamin D. Zinszer
Patrick Clark

Bibliography

Ameel, E., Malt, B. C., Storms, G., and Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language*, 60(2):270–290.

Ameel, E., Storms, G., Malt, B. C., and Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, 53(1):60–80.

Birdsong, D. and Molis, M. (2001). On the Evidence for Maturational Constraints in Second-Language Acquisition. *Journal of Memory and Language*, 44(2):235–249.

De Bot, K., Lowie, W., and Verspoor, M. (2007). A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10(01):7.

De Groot, A. M. (2012). Bilingualism and cognition. In Chapelle, C. A., editor, *Encyclopedia of Applied Linguistics*. American Psychological Association.

De Leeuw, E., Schmid, M. S., and Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition*, 13(1):33–40.

Dong, Y., Gui, S., and MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism*, 8(03):221.

Dunn, L. and Dunn, D. (2007). *PPVT-4: Peabody picture vocabulary test*. Minneapolis, MN.

Elman, J. L., Karmiloff-Smith, A., Bates, E. A., Johnson, M. H., Parisi, D., and Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. The MIT Press, Cambridge.

Ervin, S. (1961). Semantic shift in bilingualism. *The American Journal of Psychology*, 74(2):233–241.

Gogate, L., Walker-Andrews, A., and Bahrick, L. (2001). The intersensory origins of word-comprehension: an ecological–dynamic systems view. *Developmental Science*, 4(1):1–18.

Graham, C. and Belnap, R. (1986). The acquisition of lexical boundaries in English by native speakers of Spanish. *IRAL-International Review of Applied Linguistics in Language Teaching*, 24(1-4):275–286.

Hebb, D. (1949). *The organization of behavior: A neuropsychological theory*. Wiley, New York.

Hernandez, A. E. and Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, 133(4):638–50.

Hutz, M. (2004). Is there a natural process of decay? A longitudinal study of language attrition. In Schmid, M. S., Köpke, B., Keijzer, M., and Weilemar, L., editors, *First language attrition: Interdisciplinary perspectives on methodological issue*, pages 189–207. John Benjamins Publishing Company, Amsterdam.

Hyltenstam, K., Bylund, E., Abrahamsson, N., and Park, H.-S. (2009). Dominant-language replacement: The case of international adoptees*. *Bilingualism: Language and Cognition*, 12(02):121.

Johnson, J. and Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21:60–99.

Kay, P. and Regier, T. (2007). Color naming universals: the case of Berinmo. *Cognition*, 102(2):289–98.

Kohonen, T. (2001). *Self Organizing Maps*. Springer, Berlin, 3rd edition.

Kroll, J. F. and Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In Gruneberg, M., Morris, P., and Sykes, R., editors, *Practical aspects of memory*, pages 389–395. John Wiley & Sons, London.

Kroll, J. F. and Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2):149–74.

Kroll, J. F. and Tokowicz, N. (2001). The Development of Conceptual Representation for Words in a Second Language. In Nichol, J., editor, *One mind, two languages: Bilingual language processing*, chapter 3, pages 49–71. Blackwell Publishers, Malden, MA.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., and Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism*, 13(3):373–81.

Landar, H., Ervin, S., and Horowitz, A. (1960). Navaho color categories. *Language*, 36(3):368–382.

Li, P. (2009). Lexical Organization and Competition in First and Second Languages: Computational and Neural Mechanisms. *Cognitive Science*, 33:629–664.

Li, P. and Farkas, I. (2002). A self-organizing connectionist model of bilingual processing. In Heredia, R. and Altarriba, J., editors, *Bilingual sentence processing*, volume 17, pages 59–85. North-Holland.

Li, P., Sepanski, S., and Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior research methods*, 38(2):202–10.

Li, P., Zhao, X., and MacWhinney, B. (2007). Dynamic Self-Organization and Early Lexical Development in Children. *Cognitive Science*, 31(4):581–612.

Linck, J. A., Kroll, J. F., and Sunderman, G. (2009). Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychological Science*, 20(12):1507–15.

MacWhinney, B. (2005). A unified model of language acquisition. In Kroll, J. F. and De Groot, A. M., editors, *Handbook of Bilingualism: Psycholinguistic Approaches*, page 49. Oxford University Press, USA.

MacWhinney, B. (2012). The logic of the Unified Model. In Gass, S. and Mackey, A., editors, *Handbook of Second Language Acquisition. New York: Routledge*. Routledge.

Malt, B. C., Sloman, S. A., and Gennari, S. P. (2003). Universality and language specificity in object naming. *Journal of Memory and Language*, 49(1):20–42.

Malt, B. C., Sloman, S. A., Gennari, S. P., Shi, M., and Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, 40(2):230–62.

Mayor, J. and Plunkett, K. (2010). A neurocomputational account of taxonomic responding and fast mapping in early word learning. *Psychological review*, 117(1).

Meara, P. (2004). Modelling vocabulary loss. *Applied Linguistics*, 25(2):137–155.

Pallier, C., Dehaene, S., Poline, J.-B., LeBihan, D., Argenti, A.-M., Dupoux, E., and Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: can a second language replace the first? *Cerebral cortex*, 13(2):155–61.

Papafragou, A. and Selimis, S. (2010). Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes*, 25(2):224–60.

Pavlenko, A. and Malt, B. C. (2011). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian–English bilinguals. *Bilingualism: Language and Cognition*, 14(01):19–45.

Potter, M., So, K., Eckardt, B., and Feldman, L. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23(1):23–38.

Regier, T. and Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in cognitive sciences*, 13(10):439–46.

Richardson, F. M. and Thomas, M. S. C. (2008). Critical periods and catastrophic interference effects in the development of self-organizing feature maps. *Developmental science*, 11(3):371–89.

Romney, A., Weller, S., and Batchelder, W. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist*, 88(2):313–38.

Rumelhart, D. and McClelland, J. (1986). *Parallel distributed processing: Explorations in the microstruc- ture of cognition*. MIT Press, Cambridge.

Schmitt, E. (2010). When boundaries are crossed: Evaluating language attrition data from two perspectives. *Bilingualism: Language and Cognition*, (13):63–72.

Stolberg, D. and Münch, A. (2010). "Die Muttersprache vergisst man nicht" – or do you? A case study in L1 attrition and its (partial) reversal. *Bilingualism: Language and Cognition*, 13(1):19–31.

Tan, L. H., Chan, A. H. D., Kay, P., Khong, P.-L., Yip, L. K. C., and Luke, K.-K. (2008). Language affects patterns of brain activation associated with perceptual decision. *Proceedings of the National Academy of Sciences of the United States of America*, 105(10):4004–9.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition*, 4(02).

Van Hell, J. G. and De Groot, A. M. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, 1(3):193–211.

Wolff, P. and Ventura, T. (2009). When Russians learn English: How the semantics of causation may change. *Bilingualism: Language and Cognition*, 12(02):153–176.

Zhao, X. and Li, P. (2010). Bilingual lexical interactions in an unsupervised neural network model. *International Journal of Bilingual Education and Bilingualism*, 13(5):505–524.

Zinszer, B. D. and Li, P. (2010a). A SOM Model of First Language Lexical Attrition. In *Cognitive Science Society 2010*.

Zinszer, B. D. and Li, P. (2010b). Cross-language lexical interaction in object naming.