The Pennsylvania State University

The Graduate School

Department of Psychology

MEANING REPRESENTATION WITHIN AND ACROSS LANGUAGES

A Thesis in

Psychology

by

Natasha Tokowicz

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

December 2000

We approve the thesis of Natasha Tokowicz.

Date of Signature

Judith F. Kroll
Professor of Psychology
Thesis Adviser
Chair of Committee

J. Toby Mordkoff
Assistant Professor of Psychology

Karen S. Quigley
Assistant Professor of Psychology

James P. Lantolf
Professor of Spanish and Applied Linguistics

Keith Crnic
Professor of Psychology
Head of the Department of Psychology

Abstract

Most models of bilingual memory assume that concepts are similar across the bilingual's two languages. However, there are cases in which conceptual overlap may not be perfect across translation equivalents. The results of Experiment 1 show that translation production generally is slower and more error-prone the more ambiguous words are. Furthermore, ambiguity was shown to interact with conceptual salience (the degree to which a word refers to a perceptible object or ‘concreteness’, and the ease with which one can think of a context for a word or ‘context availabilty’) and cognate status (lexical overlap of translation equivalents). The effects of multiple meanings were reduced for cognates and enhanced for words high in conceptual salience, whereas the effects of multiple translations were relatively constant across word types. Furthermore, ambiguity was shown to influence performance on a within-language task, lexical decision, such that ambiguous words were responded to more quickly than unambiguous words. Conceptual salience also influenced lexical decision latencies such that words high in conceptual salience were responded to more quickly than words low in conceptual salience. Thus, ambiguity and conceptual salience do not interact in the within-language lexical decision task, and the two factors have direct and independent influences on word recognition. The results of the third experiment showed that alternate conceptual candidates become activated during the translation process, similar to results for within-language picture naming (Peterson & Savoy, 1998). The results are discussed in relation to current models of monolingual and bilingual language representation and processing.

Table of Contents

List of Figures ..... **ix**

List of Tables ..... **xii**

Acknowledgments ..... **xvi**

Chapter 1 General Introduction **1**

SHARED SEMANTIC REPRESENTATIONS ACROSS LANGUAGES 2

PARTIALLY SHARED SEMANTIC REPRESENTATIONS 4

THE PRESENT STUDY 14

GENERAL METHOD 23

Bilingual Participants *23*

Number of Translations Norms *25*

Cognate Norms *28*

Data Analysis *29*

Chapter 2 Experiment 1: Translation Production in Dutch and English **31**

HYPOTHESES 31

Direction of Translation Effects *32*

Effects of Concreteness and Context Availability (Conceptual Salience) *35*

Ambiguity Effects *37*

Interaction between Conceptual Salience and Ambiguity *38*

Cognate Effects *38*

Interaction between Conceptual Salience, Ambiguity, and Cognate Status.................. *39*

METHOD 40

Participants... *40*

Stimuli... *40*

Procedure... *42*

Data Trimming... *43*

RESULTS AND DISCUSSION... 44

Preliminary Analyses.. *45*

Effects of Conceptual Salience.. *51*

Ambiguity Effects.. *53*

Interaction between Conceptual Salience and Ambiguity.................................. *57*

Cognate Effects *61*

Interaction between Conceptual Salience, Number of Translations, and Cognate Status.. *63*

Direction of Translation Effects.. *68*

CONCLUSION 73

A Revised Model of Conceptual Salience and Ambiguity Effects *75*

Chapter 3 Experiment 2: Within-language Lexical Decision 81

WITHIN-LANGUAGE AMBIGUITY EFFECTS... 81

WITHIN-LANGUAGE CONCEPTUAL SALIENCE EFFECTS 84

METHOD 86

Participants... *86*

Stimuli *86*

Lexical Decision Procedure *88*

Results 89

Data Trimming *89*

Preliminary Analyses *90*

Reaction Time and Accuracy Data *92*

Discussion 95

Chapter 4 Experiment 3: Time Course of Translation Production 99

Time Course Task 100

Translation Variant of the Time Course Task *103*

Hypotheses 106

Method 109

Participants *109*

Design *111*

Stimuli *115*

Procedure *116*

Results 121

Data Trimming *121*

Reaction Time Analyses *122*

Accuracy Analyses *130*

Discussion 135

Methodological Concerns *138*

Conclusion 140

Chapter 5 General Discussion 141

Do Bilinguals Access the Same Meaning for Translation Equivalents? 142

Modeling Culturally-Distinct Meanings Across Languages 144

Implications for Models of Cross-Language Processing 148

Implications for the Study of Bilingual Language Processing 150

Implications for Models of Within-Language Processing 151

Future Directions 152

References 154

Appendix A Language History Questionnaire 164

Appendix B Language History Questionnaire Data from the Participants in all Bilingual Tasks 173

Appendix C Instructions for Cognate Ratings (Adapted from De Groot & Nas, 1991) 174

Appendix D Scatterplots of the Predicted Values against Residuals from Experiment 1 175

Appendix E Instructions for Concreteness Ratings (Adapted from Spreen & Schulz, 1966) 179

Appendix F Instructions for Context Availability Ratings (Adapted from Schwanenflugel, Harnishfeger, & Stowe, 1988) 180

Appendix G Scatterplot of the Predicted Values against Residuals from Experiment 2 181

List of Figures

Figure 1. The distributed feature model of bilingual memory representation (adapted from De Groot, 1992). 5

Figure 2. The distributed representation model (adapted from Van Hell & De Groot, 1998a). 9

Figure 3. The model of ambiguity effects in bilingual translation (adapted from Tokowicz & Kroll, submitted). 17

Figure 4. Estimated translation latencies (ms) as a function of conceptual salience, number of meanings, and number of forms for the meaning selected. 59

Figure 5. Estimated translation latencies (ms) at high and low levels of conceptual salience, number of translations, and cognate ratings The top panel shows translations low in cognate ratings (noncognates); the bottom panel shows translations high in cognate ratings (cognates). 66

Figure 6. Translation latencies from a post-hoc analysis of Experiment 1 data as a function of the direction of translation performed first (top panel) and lexical decision latencies from Kroll and Merves (1986) as a function of word type performed first (bottom panel). 72

Figure 7. Revised model of conceptual salience and ambiguity effects for words with one meaning and one translation. The top panel shows words high in conceptual salience and the bottom panel shows words low in conceptual salience. 76

Figure 8. Revised model of conceptual salience and ambiguity effects for words with one meaning and multiple forms. The top panel shows words high in conceptual salience and the bottom panel shows words low in conceptual salience. 77

Figure 9. Revised model of conceptual salience and ambiguity effects for words with multiple meanings and one form. The top panel shows words high in conceptual salience and the bottom panel shows words low in conceptual salience. 78

Figure 10. Graphic illustration of the task used in Experiment 3 (phonological condition). 104

Figure 11. Predicted activation functions from a serial two-stage processing model. The top graph represents a word with multiple form translations and the bottom graph represents a word with multiple meaning translations. 108

Figure 12. Scatterplot of predicted values against residuals from the direction of translation analysis (taken from participants 1, 2, and 3 in the top, middle, and bottom panels, respectively). 175

Figure 13. Scatterplots of predicted values against residuals for the two directions of translation from the conceptual salience analysis (taken from participants 1, 2, and 3 in the top, middle, and bottom panels, respectively)..176

Figure 14. Scatterplots of predicted values against residuals for the two directions of translation from the ambiguity analysis (taken from participants 1, 2, and 3 in the top, middle, and bottom panels, respectively)..177

Figure 15. Scatterplots of predicted values against residuals for the two directions of translation from the cognate analysis (taken from participants 1, 2, and 3 in the top, middle, and bottom panels, respectively)..178

List of Tables

TABLE 1 *REACTION TIME (MS) AND ACCURACY DATA (%) FROM TOKOWICZ AND KROLL (SUBMITTED; EXPERIMENT 4)* 16

TABLE 2 *LANGUAGE HISTORY QUESTIONNAIRE DATA FROM THE PARTICIPANTS OF THE BILINGUAL NORMING STUDIES BY TASK* 26

TABLE 3 *NUMBER OF TRANSLATIONS DATA FOR THE CRITICAL STIMULI OF EXPERIMENTS 1 AND 3 BY DIRECTION OF TRANSLATION* 27

TABLE 4 *MEAN COGNATE RATINGS FOR THE CRITICAL TRANSLATION PAIRS USED IN EXPERIMENTS 1 AND 3 BY DIRECTION OF TRANSLATION* 29

TABLE 5 *LANGUAGE HISTORY QUESTIONNAIRE DATA FROM THE PARTICIPANTS OF EXPERIMENT 1* 41

TABLE 6 *PROPERTIES OF THE CRITICAL STIMULI FROM EXPERIMENT 1 BY LANGUAGE OF STIMULUS* 42

TABLE 7 *AVERAGE INTERCORRELATIONS BETWEEN FACTORS FROM DUTCH TO ENGLISH AND ENGLISH TO DUTCH TRANSLATION* 47

TABLE 8 *THE AVERAGE NUMBER OF WORDS TRANSLATED CORRECTLY BY THE FIRST THREE PARTICIPANTS BY NUMBER OF MEANINGS AND NUMBER OF FORMS FOR THE MEANING SELECTED* 48

TABLE 9 *AVERAGE CHANGE IN R^2 ACROSS PARTICIPANTS DUE TO THE INCLUSION OF EACH PREDICTOR VARIABLE* 50

TABLE 10 *OVERVIEW OF THE HIERARCHICAL REGRESSION ANALYSIS TESTING FOR CONCEPTUAL SALIENCE EFFECTS* 51

TABLE 11 *OVERVIEW OF THE HIERARCHICAL REGRESSION ANALYSIS TESTING FOR AMBIGUITY EFFECTS* 54

TABLE 12 *ESTIMATED REACTION TIMES (MS) AND ACCURACY VALUES (%) AS A FUNCTION OF NUMBER OF MEANINGS AND NUMBER OF FORMS FOR THE MEANING SELECTED* 55

TABLE 13 *OVERVIEW OF THE HIERARCHICAL REGRESSION ANALYSIS TESTING FOR AN INTERACTION BETWEEN NUMBER OF TRANSLATIONS AND CONCEPTUAL SALIENCE* 58

TABLE 14 *ESTIMATED ACCURACY SCORES (%) AS A FUNCTION OF CONCEPTUAL SALIENCE, NUMBER OF MEANINGS, AND NUMBER OF FORMS FOR THE MEANING SELECTED* 61

TABLE 15 *OVERVIEW OF THE HIERARCHICAL REGRESSION ANALYSIS TESTING FOR COGNATE EFFECTS* 62

TABLE 16 *OVERVIEW OF THE HIERARCHICAL REGRESSION ANALYSIS TESTING FOR AN INTERACTION BETWEEN COGNATE STATUS, CONCEPTUAL SALIENCE, AND NUMBER OF TRANSLATIONS* 64

TABLE 17 *ESTIMATED ACCURACY SCORES (%) AS A FUNCTION OF COGNATE RATING, CONCEPTUAL SALIENCE, NUMBER OF MEANINGS, AND NUMBER OF FORMS FOR THE MEANING SELECTED* 68

TABLE 18 *OVERVIEW OF THE HIERARCHICAL REGRESSION ANALYSIS TESTING FOR DIRECTION OF TRANSLATION EFFECTS* 69

TABLE 19 *SAMPLE STIMULI FROM EXPERIMENT 2* 87

TABLE 20 *INTERCORRELATIONS BETWEEN FACTORS FOR THE REACTION TIME ANALYSIS OF EXPERIMENT 2* 91

TABLE 21 *OVERVIEW OF THE HIERARCHICAL REGRESSION ANALYSIS FOR VARIABLES PREDICTING REACTION TIME IN THE LEXICAL DECISION TASK* 93

TABLE 22 *ESTIMATED REACTION TIMES AS A FUNCTION OF NUMBER OF MEANINGS AND CONCEPTUAL SALIENCE* 94

TABLE 23 *SAMPLE STIMULI FROM THE PICTURE NAMING TASK USED BY PETERSON AND SAVOY (1998)* 101

TABLE 24 *SAMPLE STIMULI FROM EXPERIMENT 3* 105

TABLE 25 *LANGUAGE HISTORY QUESTIONNAIRE DATA FROM THE PARTICIPANTS OF EXPERIMENT 3 BY CONDITION* 110

TABLE 26 *THE NUMBER OF STIMULI IN THE PHONOLOGICAL CONDITION OF EXPERIMENT 3: PRODUCTION IN ENGLISH* 112

TABLE 27 *THE NUMBER OF STIMULI IN THE PHONOLOGICAL CONDITION OF EXPERIMENT 3: PRODUCTION IN DUTCH* 113

TABLE 28 *THE NUMBER OF STIMULI IN THE SEMANTIC CONDITION OF EXPERIMENT 3: PRODUCTION IN ENGLISH* 114

TABLE 29 *PROPERTIES OF THE CRITICAL STIMULI FROM EXPERIMENT 3: PHONOLOGICAL CONDITION: ENGLISH PRODUCTION* 117

TABLE 30 *PROPERTIES OF THE CRITICAL STIMULI FROM EXPERIMENT 3: PHONOLOGICAL CONDITION: DUTCH PRODUCTION* 118

TABLE 31 *PROPERTIES OF THE CRITICAL STIMULI FROM EXPERIMENT 3: SEMANTIC CONDITION, ENGLISH PRODUCTION* 119

TABLE 32 *REACTION TIME DATA FROM EXPERIMENT 3 BY NUMBER OF TRANSLATIONS AND CONDITION* 123

TABLE 33 *REACTION TIME DATA FROM EXPERIMENT 3 BY TYPE OF MULTIPLE TRANSLATIONS AND CONDITION* 125

TABLE 34 *TRANSLATION LATENCIES AS A FUNCTION OF SOA AND CONDITION* 129

TABLE 35 *ACCURACY DATA (%) FOR THE TRANSLATION AND NAMING TRIALS OF EXPERIMENT 3 BY CONDITION* 130

TABLE 36 *ACCURACY DATA (%) FOR WORDS WITH ONE AND MORE THAN ONE TRANSLATION OF EXPERIMENT 3 BY CONDITION* 132

TABLE 37 *ACCURACY DATA (%) FOR WORDS IN THE MULTIPLE FORM AND MULTIPLE MEANINGS CONDITIONS OF EXPERIMENT 3* 134

Acknowledgments

First and foremost, I would like to thank my mentor Judy Kroll for making the past five years livable. She was a mentor in every sense of the word, and supported me throughout this entire process. For that I am eternally grateful.

Portions of this research were funded by a teaching release award from the Research and Graduate Studies Office and a dissertation support award from The Department of Psychology at Penn State to myself, and NSF Grant BCS-9905850 to Judith F. Kroll at The Pennsylvania State University, and by a sabbatical grant to Judith F. Kroll from the Nijmegen Institute for Cognition and Information, The Netherlands.

Many people have helped to make this dissertation possible and the past five years of my life much better than they would have been otherwise. The members of the Cognition and Language Lab were great friends and helped to clarify my thinking along the way. I thank Michele Diaz, Natasha Miller, Ana Schwartz, Gretchen Sunderman, and Erica Michael who was always there in spirit if not in body. Thank you also to Natasha M. for being such a terrific namesake!

I would also like to thank the members of my committee who have helped me enormously throughout this process and have made the final version of this dissertation worlds better than the first. I thank Karen Quigley for her interest in my work, her constructive comments, her encouragement, and for keeping me on track throughout the job search process. Thank you to Toby Mordkoff for *tons* of statistical assistance and critical comments that have greatly improved this dissertation. Thanks to Jim Lantolf for agreeing to join the group at the last minute and for sharing the linguistic perspective on these issues with me.

I would like to thank all of the people who helped me during my trips to Holland. I thank Ton Dijkstra for making the arrangements for my trip and assisting me with office and lab space. I thank Niels Janssen for assisting me with housing and for introducing me to the "Hombres Nimegas" (Doron, Fonz, Joost, Kuan, Normal Bas, Ronnie, and White Bas) who entertained me during my stay. I would especially like to thank Basjo de Weerd for nearly single-handedly supplying me with all of the subjects I needed by convincing his friends and entire family to participate in my experiments. Thank you to Ana (Paloma) Blancas for being a wonderful housemate and very fun person to communicate with in both English and Spanish. I thank Ellen Janssen for being such a terrific friend, for being a great tour guide, and for encouraging (and putting up with) me throughout the job search process.

Several people were also extremely supportive in terms of research as well as social life. I thank Béryl Schulpen and Bart Hilberink for being such generous friends. They welcomed me into their home, loaned me many household (and research-related) items, and made my time in Holland very special. Janet Van Hell and her family also made both visits very special by making certain I had everything I could possibly need and by being so welcoming. Karen Mansfield provided invaluable friendship and research assistance, and was a very entertaining office mate. I thank Roel Kerkhofs for spending hours running subjects and coding tapes without ever acknowledging how boring these tasks really were. I thank Marc Rensen for being a very dedicated data analyst, for sharing his ideas about bilingualism with me, and for teaching me a thing or two about the Dutch language. I thank Walter van Heuven for discussing his research and sharing his modeling ideas with me.

Many people in the Penn State lab also helped to make this project possible. I thank Tessa Gatmaitan and Heather Shrigley for running the lexical decision experiment and collecting and coding the number of meanings norms. Thank you to Jonathan Hoover, Cassie Konkle, and Jen Scheuchenzuber for collecting and entering the ratings on the Dutch stimulus materials.

I thank Annette De Groot for making her stimulus materials and norms available for me to use in Experiments 1 and 3. I also thank Ton Dijkstra for making his materials available for use in my experiments.

To all of the wonderful women in my life for being so encouraging and setting such extraordinary examples with their own lives—you have all set your mind to accomplish something and reached far beyond your goals. For this and your friendship I thank you. And, especially to Aimee for always being my prime example setter. I think of you when I feel like I "just can't" any more and that always kicks me back into gear. Near or far, you are such an inspiration.

To Liana Brown who has become such a special friend throughout these last few years, despite having to live with me and bear my griping at all hours. Thank you for being such a wonderful editor, housemate, workout partner, travel buddy, friend, and confidant. I also thank Cathy Elsinger for letting me become part of her family during our many years together at Penn State. You're the world's best neighbor!

To Pete for his undying love and support throughout the good and bad times. We've finally done it, *together*!

To my wonderful family for never spoiling me (until now, perhaps) and for always encouraging me. You have made this all possible.

I dedicate this thesis to my Mom and Dad with much love and admiration.

Chapter 1

General Introduction

Bilinguals often claim that there are no true “translation equivalents” because words that are considered translations are frequently used in different contexts across languages and/or cultures, and therefore have somewhat different meanings in the two languages. For example, the English word “family” typically refers to the nuclear family whereas its translation in Spanish, “familia”, refers to extended family members as well. Furthermore, there are cases in which one-to-one correspondences across languages do not exist. Because different languages have some unique lexical concepts, speakers must express themselves in terms of the concepts that exist in each language (e.g., Whorf, 1956; see Green, 1998a, Lucy, 1997, and Pavlenko, 1999, for recent discussions of the implications of language-specific lexical concepts for bilingual performance). For instance, the English concept of “sibling” does not correspond to a single lexical entry in the Dutch language, and therefore one must specify 'brother or sister' to indicate “sibling” in Dutch. In addition, some translation equivalents are more broadly applied in one language than another. For example, the English word “glass” translates to both “vidrio” and “vaso” in Spanish, where the former refers to the material, and the latter refers to the vessel used for drinking. In such cases, the bilingual must not only learn translation equivalents, but also the contexts that are appropriate for their use in the two languages. This problem is particularly difficult when multiple words in the second language (L2) are used to express concepts that have the same label in the first language (L1) because it is necessary to restructure the underlying concept (MacWhinney, in press).

Despite these cross-language distinctions in meaning, most psycholinguistic models of bilingual memory representation assume that translation equivalents access the same semantic representations (e.g., Costa, Miozzo, & Caramazza, 1999; Francis, 1999b; Kroll & Stewart, 1994; Potter, So, Von Eckhart, & Feldman, 1984, but see De Groot, 1992; Van Hell & De Groot, 1998a for exceptions). Instead of emphasizing distinctions in meaning across languages, these models focus on the interconnections among the words and meanings in the two languages. They assume that bilinguals[1] are generally slower to perform tasks such as picture naming and categorization in L2 because of quantitative differences in the relative accessibility of semantic information for the two languages, but not because of a qualitative difference in the nature of the information that is retrieved (e.g., Dufour & Kroll, 1995; McElree, Jia, & Litvak, 2000). This dissertation questions these assumptions by examining the conditions under which distinctions in meaning may affect language processing.

Shared Semantic Representations Across Languages

Evidence for the common semantic hypothesis comes from a variety of sources. In cross-language priming tasks, there are reliable priming effects from semantically-related primes to target words in the bilingual's other language (Altarriba, 1990; Chen & Ng, 1989; De Groot & Nas, 1991; Frenck & Pynte, 1987; Keatley, Spinks, & De Gelder, 1994; Kirsner,

---

[1] I will use the term bilingual to refer to a person who has acquired some level of proficiency in a nonnative language but did not acquire the two languages simultaneously. Further, the bilinguals who participated in the present study were unbalanced bilinguals who were dominant in their first language.

Smith, Lockhart, King, & Jain, 1984; Meyer & Ruddy, 1974; Schwanenflugel & Rey, 1986; Tzelgov & Henik, 1990). If words in the two languages access fundamentally different semantic representations, then under the temporal constraints of the priming paradigm, no cross-language priming would be expected. The fact that it is obtained under these conditions suggests that the representation that is accessed is common to both languages.

Even more compelling evidence for the common semantic hypothesis comes from studies of translation priming, in which a prime word is masked so that participants are not aware of its identity; this prime is followed by its translation equivalent in the other language. Responses to words following their translations are faster than responses following unrelated words, even when the bilingual's two languages do not share the same alphabet (e.g., Gollan, Forster, & Frost, 1997; Jiang, 1999). These findings provide striking support for the claim that words in the two languages, even when they assume distinct lexical forms, access common semantic representations.

In the domain of production, there is additional support for shared semantics across the bilingual's two languages. In the original color Stroop task (e.g., Dyer, 1971; Mägiste, 1984; Preston & Lambert, 1969), in the translation variant of the Stroop task (La Heij et al., 1990; Miller, 1997), and in the bilingual version of the picture-word interference task (e.g., Costa et al., 1999; Hermans, Bongaerts, De Bot, & Schreuder, 1998), there are reliable cross-language interference effects when distractors in one language are semantically related to words to be spoken in the other language. This interference effect does not depend on the language of production as the magnitude of the effect is similar for words in both languages (Costa et al., 1999). This finding is easily explained by the common semantic hypothesis. If the two semantic representations were separate, then cross-language effects would require a translation process to occur. The pattern of results in Stroop-type interference experiments suggests that this is not the case.

In sum, the assumption that translation equivalents access the same meaning representations is supported by results showing effects of cross-language semantic priming, translation priming, and cross-language semantic interference. An alternative interpretation of the evidence for shared semantics is that these tasks force decisions to be made at a level at which subtle differences in meaning across languages are minimized. Indeed, few language differences are observed when tasks require high-level problem solving (e.g., Francis, 1999a, and see Francis, 1999b for a more general review of this literature). Although words in the bilingual's two languages may not access separate semantic representations, it is still possible that the meanings accessed by the two languages differ, if only slightly. As a result, many of the processing results described earlier can be accounted for by a model that assumes that there is generally a high degree of shared meaning for words identified as translation “equivalents” but that the correspondences are not exact.

Partially Shared Semantic Representations

One model that addresses the degree of cross-language meaning overlap is the distributed feature model proposed by De Groot (1992; see Figure 1). According to this model, word meanings are represented as sets of distributed features, and the meaning overlap depends on the type of word being represented. Two distinctions are made: concrete

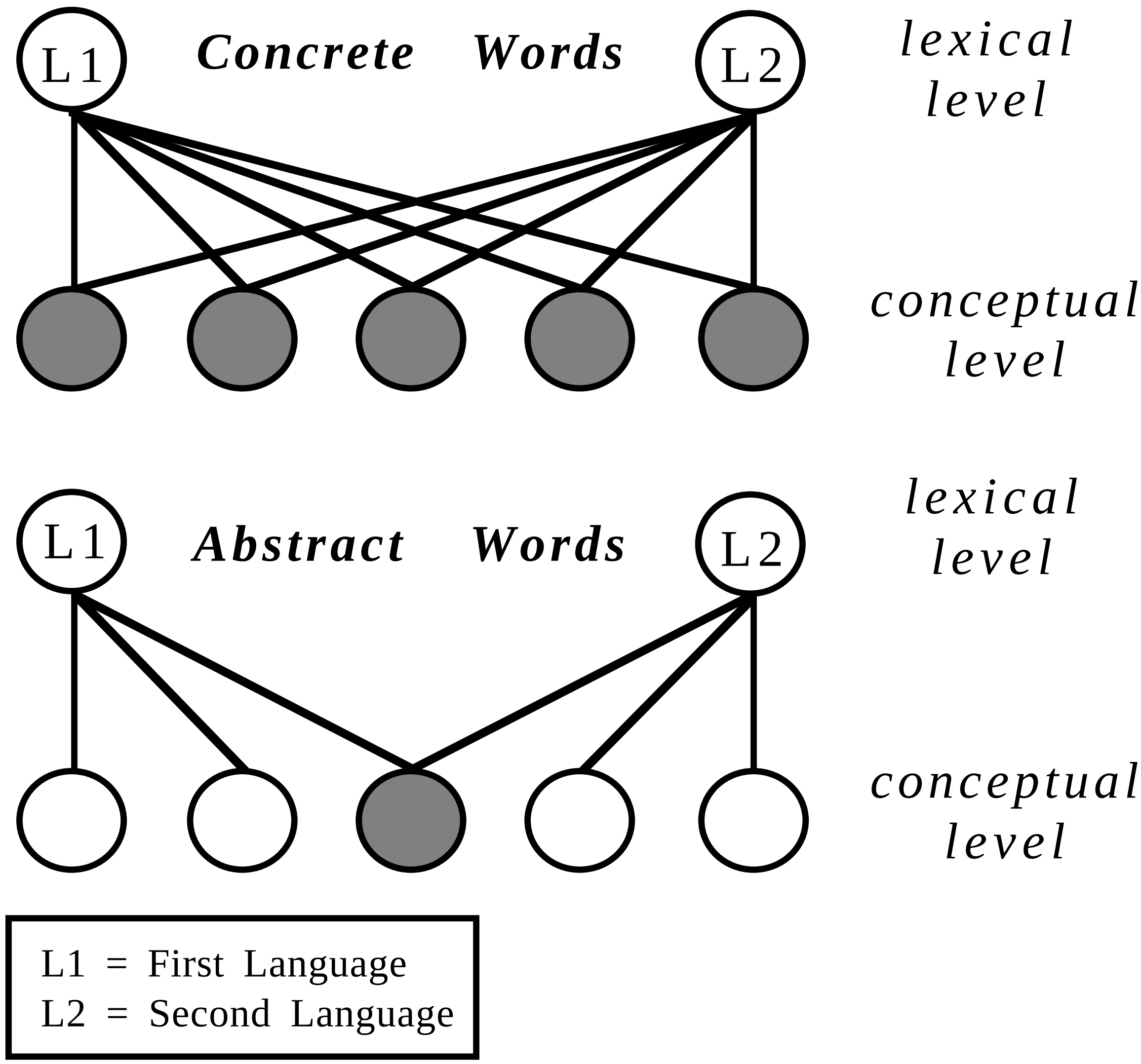


Figure 1. The distributed feature model of bilingual memory representation (adapted from De Groot, 1992).

translations differ from abstract translations and cognate translations differ from noncognate translations.

Concrete words, those that refer to perceptible entities (e.g., “table”), are assumed to have more similar meaning (indicated by more complete feature overlap) across languages than abstract words, those words that do not refer to perceptible entities (e.g., “freedom”).[2] The more complete meaning overlap for concrete words is based on the assumption that the objects to which concrete words refer have similar functions across languages and cultures. Usually, the outward appearance of, and behaviors elicited by, such objects are dictated by their functions and therefore should be similar across cultures (e.g., pens are intended for writing and their form indicates their function in most cases). In contrast, because abstract words tend to be used in different contexts across languages and cultures, they will be less similar in meaning across languages.[3]

Similar to the distinction between concrete and abstract words, De Groot (1992) also distinguished cognate and noncognate translations. Cognates are translations that share lexical form, whereas noncognates are translations that do not share lexical form; because they are translations, cognates, by definition, share meaning. False cognates or “false friends” are not translations but rather words in the two languages that share lexical, but not semantic, features (e.g., the Spanish word “red” means “net” in English, not “red”). The distributed feature model proposes that cognates share more conceptual features than noncognates. This

---

[2] Because the difference between concrete and abstract words relates to their meanings, concreteness is considered a semantic variable.

[3] Although there may be exceptions to this distinction, the consequences of such exceptions for the model were not discussed by De Groot (1992).

is because cognates often share a linguistic derivation and because bilinguals may assume that words that share lexical features will also share meaning.

A further implication of the distributed feature model is that the overlap of meaning representations across languages should reflect the manner in which the bilingual has acquired his or her second language. According to this model, adult bilinguals who have acquired their L2 in the classroom are likely to assume pure meaning equivalence across languages, regardless of word concreteness, because they are unlikely to have been exposed to cultural distinctions in meaning. This assumption is supported by evidence showing that adult bilinguals initially transfer knowledge of their L1 to their L2 (Hancin-Bhatt & Nagy, 1994; MacWhinney, 1997). However, once long-term exposure to the L2 culture has taken place, lexical reorganization may occur, and the bilingual's semantic representation of an L2 concept will come to approximate that of native speakers (Grabois, 1999). As a result of the time needed to acquire cross-language distinctions in meaning, the concrete/abstract word distinction is relevant only for bilinguals who have achieved a near-native level of proficiency or who have had L2 immersion experience.

In order to derive predictions for translation tasks from the distributed feature model, two additional assumptions have been made. First, it is assumed that higher feature overlap leads to faster processing time because the greater the number of conceptual elements shared by a pair of translations, the more activation will spread from the lexical node of a word to the node of its translation, and, therefore, the more available the translation response will be (De Groot, 1992, p. 1016). Second, the assumption is made that translation is mediated by the activation of these conceptual features (p. 1016). With these assumptions, the distributed feature model predicts that concrete words and cognates will be translated more quickly than abstract words and noncognates, respectively, because feature overlap across languages is more complete in these cases. This pattern of results has been obtained in many experiments for cognates and noncognates (e.g., De Groot, 1992; Sánchez-Casas, Davis, & Garcia-Albea, 1992), and for concrete and abstract words (e.g., De Groot, 1992; De Groot, Dannenburg, & Van Hell, 1994; Van Hell & De Groot, 1998b). In particular, De Groot (1992) showed that bilingual participants responded more quickly to concrete than abstract words on (a) *normal translation production* from L1 to L2, which involves presenting a single word for translation into the other language, (b) *cued translation production* from L1 to L2, which involves presenting a single word for translation and supplying the first letter of the expected response, and (c) *translation recognition*, which involves presenting two possible translation equivalents so that the participant may decide whether they are correct translations of each other.

Van Hell and De Groot (1998a) adapted the distributed feature model to account for feature overlap at both the lexical and conceptual levels (the distributed representation model; see Figure 2). This revised model is a verbal connectionist-type model (e.g., Hinton, McClelland, & Rumelhart, 1986), that has not yet been implemented. According to this model, words are represented by patterns of activation across an entire network of units or features. Words that have similar orthography and/or phonology share lexical features, whereas words that have similar meanings share conceptual features. Thus, cognates overlap at both the lexical and conceptual levels, noncognates overlap only at the conceptual level, and false cognates overlap only at the lexical level.

An example of this overlap is provided in Figure 2. Here, a Dutch word is given with

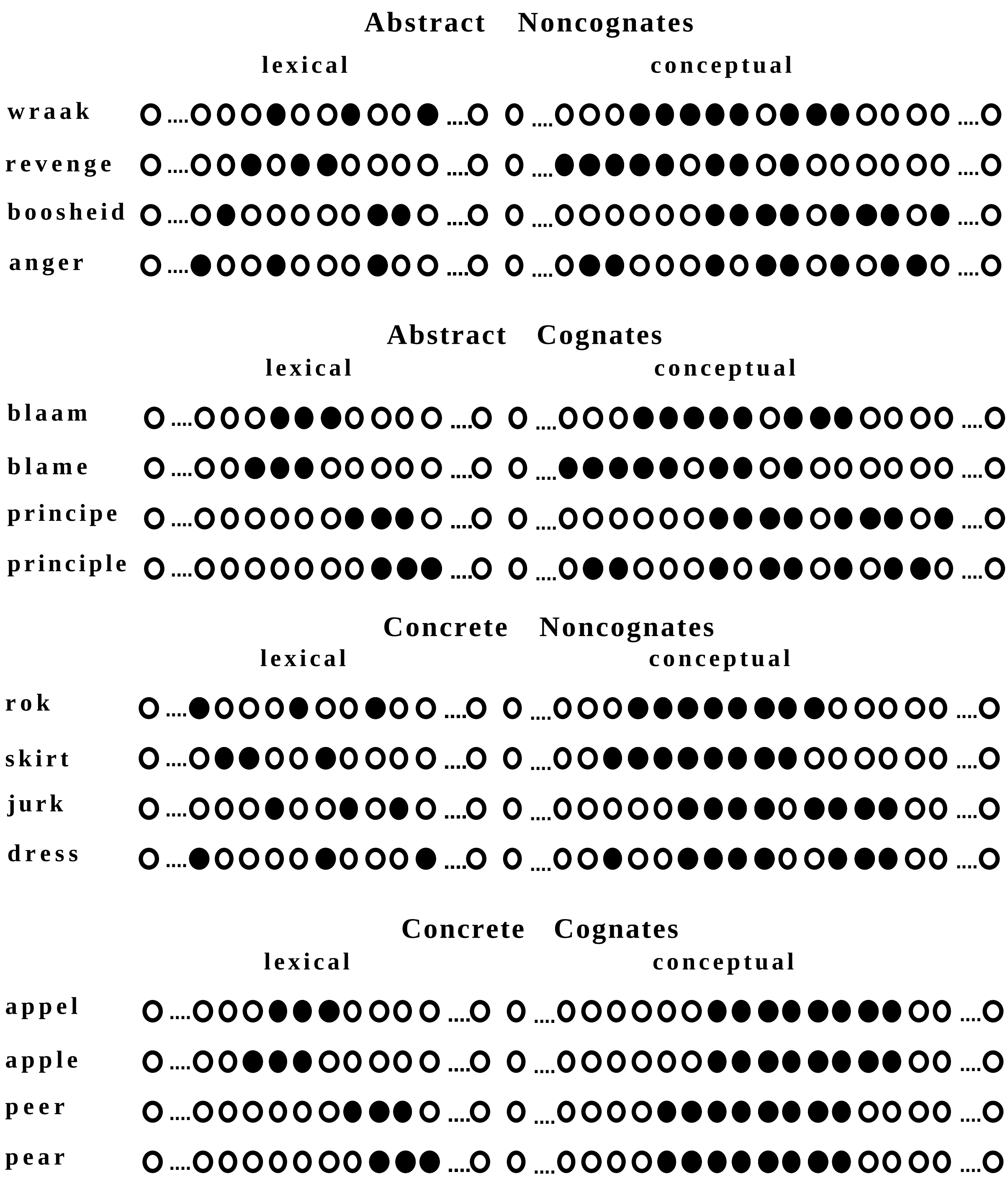


Figure 2. The distributed representation model (adapted from Van Hell & De Groot, 1998a).

its English translation below it. The first group of words are abstract noncognates, the second group are abstract cognates; the third group are concrete noncognates, and the fourth group are concrete cognates. The darkened circles represent active features, and the empty circles represent inactive features. The concrete words have more conceptual overlap with their translations than the abstract words, and the cognates have more lexical overlap with their translations than noncognates.

Although the distributed feature and distributed representation models can account for cross-language concreteness effects in translation, to be considered a general model of language representation it would be necessary to show how a model can explain concreteness effects on a within-language level as well (e.g., Kroll & Merves, 1986; Paivio, 1971; Schwanenflugel, Harnishfeger, & Stowe, 1988; Schwanenflugel & Shoben, 1983). This explanation of within-language effects was accomplished by assuming that, within a language, concrete words share more conceptual features with words to which they are semantically related than do abstract words. This is because semantically-related concrete words tend to share perceptual features, and therefore are more semantically similar than semantically-related abstract words.[4] For example, dog and cat are semantically-related concrete concepts and both have features such as [is a pet], [has fur], [has a tail], and [has four paws]. Furthermore, words that share features will partially activate one another. This difference should result in faster processing of concrete than abstract words on tasks in which

---

[4] Note that the same may not be true of associatively-related words such as "bone" and "dog".

this overlap is hypothesized to speed processing, such as word association.[5] Word associations to concrete words are performed more quickly and more consistently across sessions than word associations to abstract words (Van Hell & De Groot, 1998a).

However, the within-language concreteness effect is not limited to word association. The concreteness effect has been reported in memory tasks such as recall (e.g., Paivio, 1971) and recognition (e.g., Begg & Paivio, 1969), and in language tasks such as reading (e.g., Holmes & Langford, 1976) and lexical decision, a task in which participants judge whether letter strings are real words (e.g., James, 1975; Kroll & Merves, 1986). It is not apparent how the distributed feature and representation models could account for concreteness effects in such memory and language tasks because it is unclear why within-language semantic overlap would aid processing in these tasks.

If the conceptual representations of concrete words are more easily accessed and remembered than those for abstract words, even within a single language, then a unitary semantic system with differentiated concepts (e.g., concrete and abstract) may be sufficient to account for concreteness effects in bilingual translation. That is, it may not be necessary to assume differential overlap among translation equivalents; any factor that slows conceptual access in general will also slow bilingual translation to the extent that conceptual access is a component of the task. According to this view, bilingual performance reflects a cross-language manifestation of distinctions among concepts. If this is true then cross-language

---

5 In this task, the stimulus will activate semantically and lexically-related words quickly because of their shared patterns of activation, and, although any related word would be a potential response, semantic features play a stronger role than lexical features in word association (Van Hell & De Groot, 1998a).

concreteness effects should be similar to within-language concreteness effects to the extent that the tasks used are comparable in the nature of semantic processing they require.

The initial explanation for concreteness effects in language and memory tasks comes from Paivio's (1971) dual-coding theory. According to this theory, concrete words are presumed to be understood or accessed more quickly than abstract words because concrete words are represented dually in memory, once in the verbal system and once in the imaginal system. Abstract words are coded only in the verbal system because they have no imaginal code associated with them.

An alternative explanation for the concreteness effect is provided by the context availability theory (e.g., Kieras, 1978; Schwanenflugel et al., 1988; Schwanenflugel & Shoben, 1983). Context availability is the extent to which one can generate a context for a given word and, in general, concrete words are higher in rated context availability than abstract words. Furthermore, a rating study showed that concrete words generally appear in fewer contexts than abstract words (Schwanenflugel & Shoben, 1983). Assuming a system with a set amount of available activation, the more concepts to which a word is associated, the less strong the association between the word and any one concept will be (e.g., the "fan effect"; Anderson, 1974, 1983). Based on these findings, it is hypothesized that concrete words are strongly connected to few nodes in memory whereas abstract words are weakly connected to many nodes in memory (Schwanenflugel & Shoben, 1983).

The context availability theory is supported by two important findings. First, the concrete-word advantage in lexical decision was shown to disappear when supportive context was provided (e.g., sentential context prior to lexical decision; Schwanenflugel & Shoben, 1983). Furthermore, concrete and abstract words that are matched for context availability are responded to equally quickly in the out-of-context lexical decision task (Schwanenflugel et al., 1988). Taken together, these results suggest that the representations of concrete and abstract words tend to differ but that this difference may be more the product of the availability of context than of concreteness *per se*.

Both the dual-coding and context availability theories provide explanations for the difference in the speed of semantic access for concrete and abstract words. These explanations can be extended to cross-language tasks that involve semantic access. In contrast, because the distributed feature model focuses on cross-language meaning overlap, it does not provide as comprehensive an account of concreteness effects as the dual-coding and context availability theories.

Although cross-language processing is influenced by factors that affect semantic processing in general, there are some additional factors that uniquely affect cross-language performance. For example, cognate status relates to the lexical and semantic overlap of words across languages. Furthermore, although there is ambiguity within a language (e.g., in terms of the number of meanings words have), number of translations is a form of ambiguity that is relevant to only cross-language processing. For example, as mentioned earlier, the word “glass” in English has two Spanish translations. When an English-Spanish bilingual translates the word “glass”, the appropriate translation in Spanish must be selected from two options, which may create competition. The effects of number of translations have not been studied extensively in the past but will be used in this study as a tool for understanding the consequences of imperfect meaning overlap across languages.

The Present Study

The present study was aimed at exploring the consequences of cross-language meaning distinctions due to ambiguity in terms of number of translations. Schönpflug (1997) showed that abstract words were more likely than concrete words to have multiple translations across languages[6], and that words with multiple translations are translated more slowly than words with one translation. These findings show that both concreteness and ambiguity are important factors in bilingual language processing.

Tokowicz and Kroll (submitted) further explored the consequences of concreteness and number of translations for translation performance. Translation production was selected as the task of interest because these effects may be exaggerated during production and because translation is the primary task in experiments examining cross-language concreteness effects. The translation task shares many features of normal language production, however the process begins at a different point. In normal language production, one begins with an intended message and then selects the forms that will be used to express that message. In the translation task, one is given a form in one language that is then translated to a word in the other language. Thus, the translation production process is initiated with a word rather than a thought as in normal language production.

---

6 From German to English, concrete and abstract words were assigned 2.31 and 4.67 translations, respectively. From English to German, concrete and abstract words were assigned 2.49 and 3.00 translations, respectively [overall analysis: F (3, 250) = 16.61, $p <$ .01; Schönpflug, 1997].

In two experiments, Tokowicz and Kroll (submitted) found that words with only one or a clearly dominant translation (as determined by a bilingual experimenter and verified with a bilingual dictionary) showed no concreteness effect in translation. These results are inconsistent with past studies in which concreteness effects in translation have been reported (e.g., De Groot, 1992; De Groot et al., 1994). In addition, the lack of a concreteness effect in translation calls existing theories of these effects into question because such theories propose the effects to be ubiquitous. Tokowicz and Kroll then obtained number of translations norms for a set of English words and their translations in Spanish. In this norming task, bilinguals wrote the first translation they thought of for each word; the total number of distinct (correct) translations across participants was then calculated. Words with more than one translation in either or both directions of translation were classified as "ambiguous" and words with only one translation in both directions of translation were classified as "unambiguous".

These norms were used to select stimuli for an experiment in which ambiguity and concreteness were manipulated. Tokowicz and Kroll found that ambiguity had an overall effect on translation, such that the words with more than one translation were translated more slowly than words with one translation.[7] Furthermore, concrete words were translated more quickly than abstract words overall in the analysis by participants. However, concreteness and number of translations interacted such that concrete words were not influenced by number of translations (see Table 1). Thus, the cross-language concreteness effect reported by De Groot and colleagues (e.g., De Groot, 1992; De Groot et al., 1994; Van Hell & De Groot, 1998b) was only observed for words with multiple translations across languages.

[7] More unambiguous than ambiguous words were included in that experiment so that the participants would not determine the purpose of the experiment.

Table 1

Reaction Time (ms) and Accuracy Data (%) from Tokowicz and Kroll (Submitted; Experiment 4)

	Number of Translations					
Direction of Translation	More than One Translation			One Translation		
	Abstract	Concrete	Concreteness Effect	Abstract	Concrete	Concreteness Effect
L1 to L2	1526 (27)	1374 (46)	152 (19)	1323 (50)	1322 (66)	1 (16)
L2 to L1	1359 (35)	1264 (52)	95 (17)	1190 (58)	1290 (70)	-100 (12)

Note. Accuracy data (%) are shown in parentheses. Negative effect sizes indicate a reversal of the typical concreteness effect (i.e., faster and more accurate responses to concrete than abstract words).

Tokowicz and Kroll proposed a model that gives an account of why the concreteness effect is present only for words with multiple translations (see Figure 3). According to this model, during translation a word activates all of its meanings in parallel (e.g., Onifer & Swinney, 1981; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979).[8] These meanings will in turn activate all of their

8 Some additional assumptions have been added to this model for the present purposes.

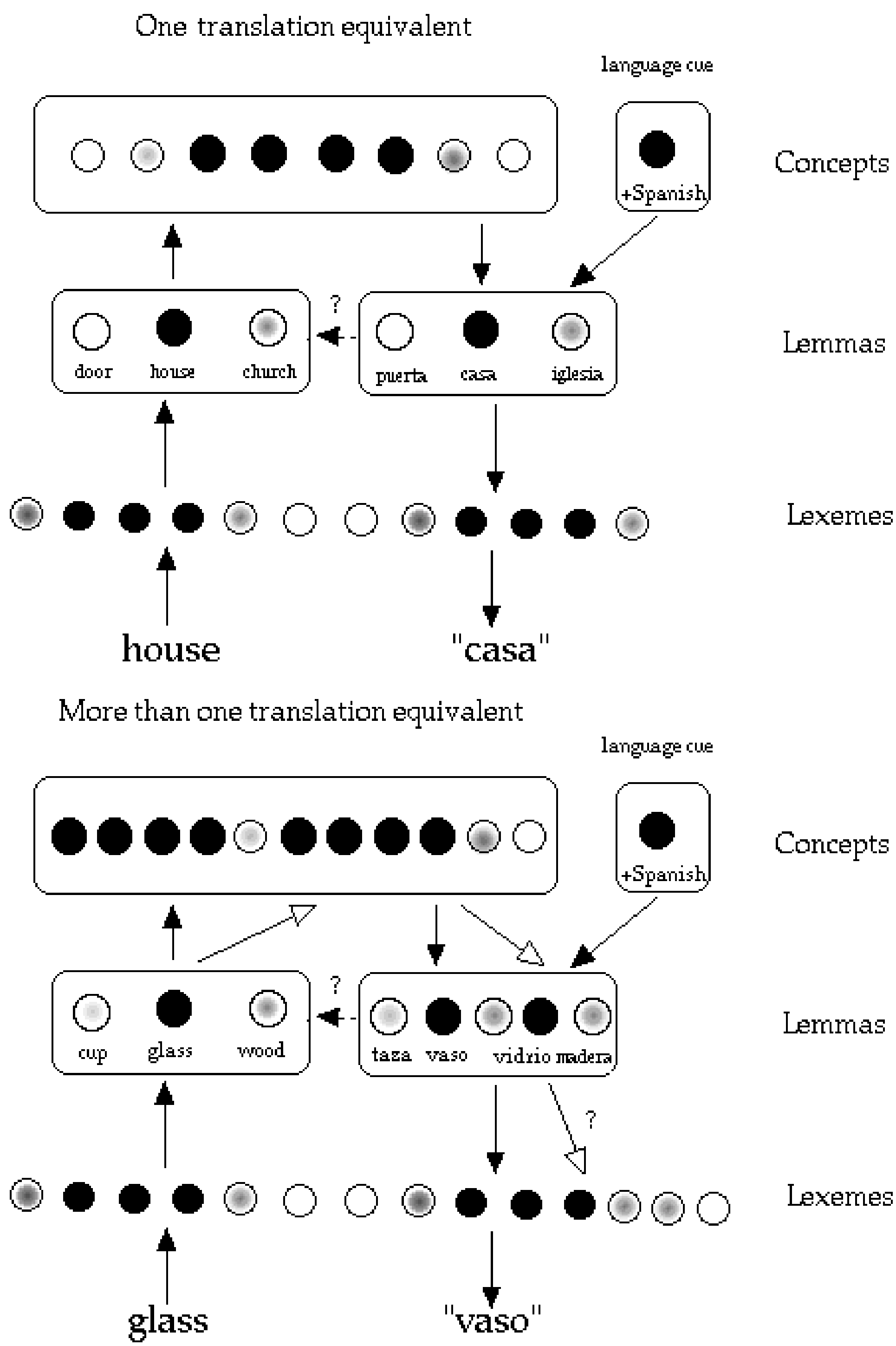


Figure 3. The model of ambiguity effects in bilingual translation (adapted from Tokowicz & Kroll, submitted).

forms (translations) in the other language.[9] As the meaning selection process proceeds, there is increased activation of the relevant forms in the other language. When the meaning selection completes fully, the form(s) connected to the selected meaning will get a large increase in activation because final resolution of one meaning will lead to inhibition of other meanings, and in turn inhibition of the translations related to those meanings. This occurs because the model assumes a system that includes lateral inhibition, which causes an active node at one level of a system to inhibit activation of other nodes at that level.

When there are multiple nodes possible for output (i.e., multiple forms), translation will not be possible until the meaning resolution is complete because there will not be enough activation to reach threshold until there is additional activation due to the completion of meaning selection and the resulting lateral inhibition. This is because there is a set amount of activation in the system and this activation must be shared among all associated nodes, so the more nodes the less activation sent to each node (e.g., Anderson, 1974, 1983). Furthermore, when multiple nodes are active, there is *active competition* among the nodes (e.g., Jacobs & Grainger, 1992; McClelland & Rumelhart, 1981, 1985). This is because the translation task requires that only one node be selected for output. The greater the number of competing nodes, the more output will be slowed.

Therefore, in the cases in which only one form is possible, a difference between concrete and abstract words will not be observed because the *initial* processes of meaning

---

[9] There is recent evidence to suggest that words in both languages are activated during within-language comprehension tasks (e.g., lexical decision, Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Van Hell & Dijkstra, 2000) and language production tasks (e.g., picture naming, Hermans et al., 1998). This is even true when the participants are unaware of the bilingual nature of the task. Therefore, it is likely that all translations of a word will become activated in the context of this translation task.

activation will result in a translation being selected. However, when there are multiple forms, concrete words will be translated more quickly than abstract words because they will get an increase in activation from the word meanings sooner than the abstract word meanings (i.e., meaning resolution will complete earlier for concrete than abstract words). That is, one meaning needs to be fully selected only when multiple forms are available. Despite the advantage for concrete over abstract words with multiple forms, concrete words with multiple forms should be translated more slowly than words with one form. This is indeed the case when one examines the absolute reaction times from the Tokowicz and Kroll study, although there was not a significant difference between the conditions, which may have been due to low statistical power.

Furthermore, words with a single form and multiple meanings will not be translated as slowly as words with multiple forms and a single meaning. That is because although the existence of multiple meanings may slow the meaning activation process somewhat, relatively fast resolution of the proper translation will be possible. Thus, it is predicted that the consequences of having multiple forms will be more severe than the consequences of having multiple meanings.

This model also allows for the possibility that there may be lexical-level connections from L2 words to their corresponding translations in L1. Therefore, translation from L2 to L1 could take place via these connections, thereby bypassing conceptual activation. This manner of translation would eliminate conceptual but not lexical-level competition. There are implications of the stimulus characteristics for this distinction between lexical and conceptual level competition. In the Tokowicz and Kroll study, concrete words were equally likely to have multiple translations due to near synonyms in one of the languages (e.g., “sofá” in

Spanish translates to both "couch" and "sofa" in English) and to multiple meanings (e.g., the drinking vessel meaning of the English word "glass" translates to "vaso" in Spanish, whereas the material meaning of "glass" translates to "vidrio"). In contrast, abstract words with multiple translations primarily had multiple translations due to multiple meanings within a language. Thus, it is possible that ambiguity itself is not equivalent for concrete and abstract words, which could have resulted in an apparent interaction between ambiguity and concreteness. If L2 to L1 translation is lexically mediated, the model Tokowicz and Kroll presented predicts that only lexical-level ambiguity (i.e., multiple forms or synonyms) and not conceptual-level ambiguity (i.e., multiple meanings) will influence translation performance in that direction.

The model and assumptions outlined here have several implications for translation performance. First, words that have both multiple meanings and multiple translations will suffer from competition at both levels of the system. This should lead to even more detrimental performance than competition at one level alone. However, there are very few words that fit into the category of words with multiple meanings and forms. Additionally, multiple forms will slow performance to a greater extent than multiple meanings. Furthermore, this model supposes that although there are differences between concrete and abstract words in the speed with which they access their meanings, these differences will not be observed in the translation task unless there are multiple nodes possible for production. Finally, there may be implications of translation direction in that L2 to L1 translation may be influenced only by lexical-level ambiguity.

The model presented by Tokowicz and Kroll was based on the limited evidence that was available on the effects of concreteness and ambiguity on translation performance. The experiments of the present study were used to test the predictions of this model. Furthermore, several aspects of the interaction between ambiguity and concreteness observed by Tokowicz and Kroll (submitted) warranted further investigation. First, a limitation of that study was that words with more than one translation in either *or* both directions of translation were considered ambiguous. However, some translation pairs had multiple translations in only one of the directions of translation. This method of classification may have over or underestimated the effects of ambiguity because some words that were classified as ambiguous were not ambiguous in the direction of translation being performed. Therefore, in this study, the ambiguity in the particular direction of translation the participant is performing will be used to predict performance.

Second, the reason that words had multiple translations was not taken into account in that experiment. As mentioned above, words with multiple synonym (form) translations create competition at the lexical level whereas words with multiple meaning translations create competition at the conceptual level. In this study, lexical and conceptual level ambiguity will be disentangled.

An additional concern is that the majority of the past studies showing a concreteness effect in translation were conducted using Dutch-English bilinguals as participants. It was important to extend the findings of Tokowicz and Kroll to this group of bilinguals because the nature of the language learning experience of the bilinguals being tested may influence the likelihood that concreteness effects are observed. The distributed feature model (De Groot, 1992) proposes that bilinguals with immersion experience or near-native proficiency in L2 are most likely to show a concreteness effect in translation. The Dutch-English bilingual participants in the experiments by De Groot and Van Hell (e.g., De Groot, 1992; De Groot, 1993; De Groot, 1995; De Groot et al., 1994; Van Hell & De Groot, 1998a, 1998b) were late second language learners who are typically dominant in their L1, Dutch, and who have most often not lived in an English-speaking environment. However, the use of English among Dutch university students is unusual because there is a high expectation for reading and speaking in English on a frequent basis and because English is used a great deal in the culture at large. These Dutch-English bilinguals therefore live and work in an environment that, although not technically an L2 immersion environment, requires relatively continuous and interchangeable use of L2. In contrast, whereas the bilinguals in the Tokowicz and Kroll study generally had L2 immersion experience, they used the two languages in distinct contexts. Thus, the two bilingual groups differ in their immersion experience and the degree to which they interchange their two languages. The results of the Tokowicz and Kroll study suggested that the immersion experience of the participants did not determine the magnitude of the concreteness effects as the distributed feature model predicts; however it is still possible that the other differences between the bilingual groups could influence the effects observed.

Finally, the Tokowicz and Kroll study did not include cognates as critical items because there were too few ambiguous cognates in their stimulus set. The present study allowed for a more detailed analysis of cognate effects in relation to those of concreteness and ambiguity.

In sum, Experiment 1 was aimed at exploring the findings of Tokowicz and Kroll further, using the translation production task, so that the independent influences of concreteness and ambiguity could be investigated. The present experiment extended the Tokowicz and Kroll study by investigating the influence of the number of translations in the particular direction of translation that the participants were performing, by examining the effects of the type of multiple translations of the words, by using a different population of bilinguals, and by including cognate translations.

Experiment 2 was an attempt to determine whether the interaction between conceptual salience and ambiguity observed across languages is due to a general property of conceptual representation or to a property of language processing that is specific to bilinguals. Finally, Experiment 3 was a first attempt at examining the time course of ambiguity effects on translation using a stimulus onset asynchrony (SOA) manipulation.

General Method

Bilingual Participants

In Experiments 1 and 3 and the bilingual norming tasks, the participants completed a language history questionnaire designed to investigate their second language learning experiences (see Appendix A for the questionnaire and Appendix B for the mean data by task). They rated their first and second language reading, writing, speaking, and oral comprehension abilities on a scale from one to ten. They also indicated the age at which they began learning their L2 and what types of exposure they had learning their L2. The participants completed the language history questionnaire at the end of the session whenever an on-line task was used.[10] The participants in the bilingual tasks all came from the same population; they were students at the University of Nijmegen, Nijmegen, The Netherlands. Dutch-English bilinguals comprise a fairly homogeneous group of bilinguals in terms of the age at which they begin to study English, and the context and length of study of the language. These bilinguals have primarily studied English in a school setting, and most have had some training in French and German as well. Although French and German are optional courses of study, English is obligatory, therefore these students have received more training in English than any other foreign language (Van Hell, 1998). Furthermore, Dutch university students are expected to both speak and read in English for their study. It was a condition of inclusion in the present study that the bilinguals have Dutch as a first language and consider English their second language.

Given the relative homogeneity of this bilingual population, it was unlikely that there would be major differences among the groups of participants in the various tasks, however the language history questionnaire data were used to verify that the groups of bilingual participants did not differ in terms of the age at which they began learning English, the amount of time they studied English, the amount of immersion experience they had with English, and their self-ratings of fluency. Although self-ratings are an imperfect measure of fluency because learners are sometimes not good at determining their level of fluency (e.g., MacIntyre, Noels, & Clement, 1997), these ratings were used as a descriptive measure of the

---

[10] The participants completed the questionnaire at the end of the session so that their performance on the on-line task was not influenced by their responses on the questionnaire. It is possible that their responses on the questionnaire were influenced by their performance on the task, however, this is less problematic because the main dependent measures were obtained during the on-line task.

groups of bilinguals along with the more objective measures (e.g., age began L2 study, number of years studied L2). Indeed, the bilingual groups only differed along two dimensions (L2 writing and speech comprehension ability), and this was due to the relatively high self-ratings of one group of participants. The major results of this study are not compromised by these differences.

No one participated in more than one experiment of this study because the same stimuli were used in several tasks. Furthermore, no stimuli were repeated in a given experiment.

Number of Translations Norms

Number of translations norms were collected for the set of Dutch and English words used in Experiments 1 and 3. These words had been used in several previous experiments (De Groot, 1992; De Groot et al., 1994; Dijkstra, Grainger, & Van Heuven. 1999; Van Hell & De Groot, 1998a, 1998b). The participants in the norming task were 24 Dutch-English bilinguals from the same population as those who participated in Experiments 1 and 3; see Table 2 for the language history questionnaire data from these participants. The words were divided into several groups and printed in random order in booklets. Each participant wrote their first spontaneous translation for each word and translated words from only one group and into only one language. The responses were coded for accuracy using Prisma Dutch-English and English-Dutch dictionaries (1990) and by a native Dutch speaker who was an advanced student of English at the University of Nijmegen. The number of correct possible

Table 2

Language History Questionnaire Data from the Participants of the Bilingual Norming Studies by Task

	Task	
Measure	Number of Translations Norms	Cognate Norms
Age (years)	21.6 (3.0)	21.2 (2.3)
Age Began L2 (years)	10.5 (2.1)	9.0 (2.5)
Time Studied L2 (years)	9.3 (3.6)	9.9 (3.8)
L2 Immersion Experience (months)	5.0 (12.3)	2.0 (4.0)
L2 Reading Ability	7.5 (1.3)	8.2 (.9)
L2 Writing Ability	6.5 (1.3)	7.1 (1.1)
L2 Conversation Ability	6.9 (1.2)	7.3 (1.3)
L2 Speech Comprehension Ability	8.0 (1.0)	8.2 (.9)

Note. Standard deviations are given in parentheses. Reading, writing, conversational, and speech comprehension ability were rated on a 10-point scale where 1 indicated the lowest level of ability and 10 indicated the highest level of ability.

translations for each word was calculated based on the total number of expected translations (those that had been considered correct in the original experiment), synonyms, other meanings (when a different meaning of the stimulus word was translated), verb meanings, and colloquial uses. The number of possible translations for each word in each direction of translation and the number of meanings to which the translations corresponded for a given word in a given direction were calculated (see Table 3). The data show that the majority of the translations given were expected responses and the overall number of translations was fairly low, suggesting that the stimulus set had a large proportion of unambiguous stimuli.

Table 3

Number of Translations Data for the Critical Stimuli of Experiments 1 and 3 by Direction of Translation

	Direction of Translation	
Measure	Dutch to English	English to Dutch
Number of Translations	1.2 (.5)	1.3 (.7)
Number of Meanings Translated	1.1 (.3)	1.1 (.4)
Percent Expected Translations	80.0 (26.7)	80.0 (30.0)
Percent Synonym Translations	5.0 (15.0)	6.7 (20.0)
Percent Other Meaning Translations	1.7 (10.0)	5.0 (15.0)
Percent Verb Translations	.1 (1.7)	.6 (3.3)
Percent Colloquial Translations	.2 (3.3)	0 (0)
Percent Incorrect Translations	13.0 (19.8)	7.7 (17.3)

Cognate Norms

Ratings of cognate status were obtained for all correct translation pairs that had been given in the number of translations norming study (i.e., not only for the original translation pair, but also for each stimulus paired with every other *possible* translation). These ratings were obtained from 16 Dutch-English bilinguals (see Table 2 on page 26 for their language history questionnaire data). Participants rated the similarity of the words in each translation pair in terms of their combined spelling and sound similarity on a 7-point scale (e.g., De Groot & Nas, 1991; see Appendix C for the rating instructions).[11] The mean cognate ratings for the critical stimuli are given in Table 4. Although the mean ratings were at the similar end of the scale, there were many items that could be considered noncognates (in both directions, at least one half of the translation pairs received mean ratings of less than 2). The relatively high average rating presumably reflects the great number of cognates in Dutch and English (approximately 50 % for common words, Van Hell, 1998).

[11] Because the pilot participants reported difficulty rating the word pairs with respect to their spelling and sound similarity without taking their meaning similarity into account, in the final version of the task, participants also rated the pairs in terms of their meaning similarity across languages. This rating was completed just prior to the cognate rating for each word pair.

Table 4

Mean Cognate Ratings for the Critical Translation Pairs used in Experiments 1 and 3 by Direction of Translation

Direction of Translation	Mean Cognate Rating
Dutch to English Translation	4.7 (.9)
English to Dutch Translation	5.0 (.9)

Note. Standard deviations are given in parentheses.

Data Analysis

In the study reported by Tokowicz and Kroll, number of translations and concreteness were orthogonally manipulated. The stimuli were selected such that the words in the various conditions did not differ with respect to word length or frequency. Therefore, in that study, an analysis of variance was used to analyze the data. Because words vary along many dimensions, it is sometimes necessary to select a sample of stimuli that does not necessarily represent the distribution of words in a natural sample to orthogonally manipulate factors. Because of this issue, the words used in the present experiment were those that had been used in previous studies and presumably represent a more natural sampling of words than were used by Tokowicz and Kroll. As a result of the stimulus selection procedure, the words vary along a number of dimensions including length and frequency. Therefore, the effects of covariates such as length and frequency were removed from the regression analyses prior to the determination of the effects of the critical predictor variables. Although the data were analyzed differently in this study and the Tokowicz and Kroll study, and finer distinctions among types of ambiguity were made in the present study, whenever possible, the data from the present study are discussed with respect to the corresponding results in the Tokowicz and Kroll study.

Chapter 2

Experiment 1: Translation Production in Dutch and English

This experiment extended the results of the Tokowicz and Kroll study by including cognate translations, and by more adequately examining the effects of multiple meanings and multiple translations on translation performance. Furthermore, Dutch-English bilinguals were tested in this study to keep the testing situation most similar to that used in past studies in which concreteness effects had been observed (e.g., De Groot, 1992; De Groot et al., 1994). Finally, this experiment tested certain predictions made by the model presented by Tokowicz and Kroll.

Hypotheses

I will now describe the hypotheses based on the models that were reviewed in the introduction. These models include models of conceptual salience (i.e., the dual coding and context availability models), the distributed models of bilingual memory representation (i.e., the distributed feature model and the distributed representation model), the revised hierarchical model, and the model presented by Tokowicz and Kroll. These models are not mutually exclusive, and each model makes predictions for only some of the analyses that were performed.

Direction of Translation Effects

Past research on bilingual memory suggests that bilinguals access meanings more readily from L1 words than from L2 words. This asymmetry in the strength of conceptual connections is assumed to be due to the manner in which L2 words are typically learned; in the classroom, L2 words are often paired with their L1 translations and therefore will be more strongly connected to their translations than to their corresponding meanings. Furthermore, for adult bilinguals who learn L2 in an immersion environment, the L1 conceptual system is already in place when the L2 is learned, therefore the connections between L1 words and meanings will be stronger than the connections between L2 words and meanings (the revised hierarchical model, Kroll & Stewart, 1994). A consequence of this asymmetry is a corresponding asymmetry in translation performance; translation from L1 to L2 takes longer to perform and is more error-prone than translation from L2 to L1 (e.g., De Groot et al., 1994; Kroll & Stewart, 1994; Sánchez-Casas et al., 1992; Sholl, Sankaranarayanan, & Kroll, 1995).

Revised Hierarchical Model. The interpretation of this asymmetry made by the revised hierarchical model (Kroll & Stewart, 1994) is that the L1 to L2 direction of translation always involves conceptual activation (which requires time to complete, and therefore increases translation time). In contrast, the L2 to L1 direction can be accomplished solely at a lexical level via word-to-word associations among translation equivalents.

The hypothesis that only L1 to L2 translation always involves conceptual mediation is also supported by the finding that semantic variables (e.g., concreteness, context availability) influence the L1 to L2 direction of translation to a greater extent than the L2 to L1 direction (De Groot et al., 1994; Van Hell & De Groot, 1998b), and the finding that there is more semantic priming from L1 primes to L2 targets than the reverse (e.g., Altarriba, 1990; Gollan et al., 1997; Keatley et al., 1994; Tzelgov & Eben-Ezra, 1992). Finally, Kroll and Stewart (1994) showed that another semantic variable, semantic categorization, only influenced L1 to L2 translation. They presented words for translation either in semantically-categorized lists (e.g., all the vehicles, then all the fruits, then all the clothing) or randomly mixed lists (e.g., tomato, car, dress, etc.). There was only an effect of presentation type for L1 to L2 translation.

Despite this evidence, there have been several failures to find an asymmetry in translation (e.g., De Groot, et al., 1994; De Groot & Poot, 1997; La Heij, Kerling, & Van der Velden, 1996). There have also been several cases in which semantic variables influenced L2 to L1 translation, although this was typically to a lesser extent than for L1 to L2 translation (e.g., De Groot et al., 1994; La Heij et al., 1996, but see De Groot & Poot, 1997, for an exception). Together, these results have been considered evidence against a strong version of the revised hierarchical model.

These failures to find a translation asymmetry are most likely due to various methodological differences among the studies. In most cases in which no difference between the two directions of translation are observed, the direction of translation is manipulated between participants (De Groot, 1992, L1 to L2 only; De Groot et al., 1994, L2 to L1 only) or different stimuli are used in the two directions of translation (De Groot & Poot, 1997). When direction of translation is manipulated between participants, the requirement to perform only one of the two directions may lead to a different strategy or threshold for the task than when participants perform both directions of translation. In addition, individual differences in the groups of participants that perform the two directions could influence the results.

Furthermore, sometimes participants are familiarized with the stimuli prior to the experiment to reduce errors or are presented the same words several times within an experimental session (e.g., Altarriba & Mathis, 1997; La Heij et al., 1996). Both of these procedures could lead to the strengthening of the connections between the L2 words and their meanings (which is believed to be the difficult aspect of L1 to L2 translation). Such strengthening would therefore decrease the difference between the two directions of translation. In support of this argument, picture naming, a task that involves the concept to word component of the translation task, primes L1 to L2 translation and eliminates the translation asymmetry (Sholl et al., 1995).[12]

Finally, several semantic factors have been used to determine whether semantic access has taken place during translation (e.g., cross-language semantic priming, semantic interference, concreteness effects, category interference). Because different factors are likely to engage different levels of semantic processing, these different tasks may be more or less likely to reveal differences between the two directions of translation.

---

[12] In that study, bilinguals first named pictures in L1 and L2. Then, the participants performed translation, and some of the concepts had been presented during the picture naming task. The transfer of L2 picture naming to L1 to L2 translation and L1 picture naming to L2 to L1 translation were compared. Picture naming only primed L1 to L2 translation, the direction of translation hypothesized by the revised hierarchical model to include the same L2 to concepts connection as picture naming. This finding suggests that the two directions of translation differ qualitatively rather than just quantitatively.

In sum, the revised hierarchical model predicts that semantic variables (e.g., concreteness) will influence L1 to L2 translation more than L2 to L1 translation. The effect of these variables on translation in the two directions will be evaluated in this experiment by comparing the two directions of translation.

Tokowicz and Kroll Model. The model presented by Tokowicz and Kroll allows for the possibility that translation from L2 to L1 can be performed at the lexical level. However, this need not be the case. Therefore, this model does not make a clear prediction regarding direction effects, however either no difference between the two directions of translation or the same asymmetry that is predicted by the revised hierarchical model should be observed.

Distributed Models. The distributed feature and representation models assume that translation is always conceptually mediated. Therefore, these models predict that conceptual variables will influence the two directions of translation to a similar extent, and that no difference between the two directions of translation will be observed.

Effects of Concreteness and Context Availability (Conceptual Salience)

Although concreteness and context availability are distinct factors, they are typically correlated (.73 and .65, Schwanenflugel et al., 1988, Experiments 2 & 3, respectively; .88, Schwanenflugel & Shoben, 1983). These factors were also highly correlated in the present study (.82 from Dutch to English and .80 from English to Dutch, see Table 7 on page 47) and therefore they will be used together to determine the influence of these factors on translation performance. The set of these two factors will be referred to as “conceptual salience”. When the individual factors are referred to, their individual names will be used.

Tokowicz and Kroll Model. The model proposed by Tokowicz and Kroll predicts that the differences between words high and low in conceptual salience will be observed only when ambiguity is present. Therefore, there may be an overall effect of conceptual salience, but this effect will be qualified by an interaction between conceptual salience and ambiguity. Furthermore, because this model assumes that translation from L2 to L1 could be accomplished at a lexical level, conceptual salience effects may be smaller (limited to multiple forms) in this direction of translation.

Distributed Models and Models of Conceptual Salience. The distributed feature and representation models and the dual-coding and context availability theories predict that words high in conceptual salience will be translated more quickly than words low in conceptual salience. This difference may be attributable to either their more complete feature overlap across languages (e.g., De Groot, 1992), or instead to a more general principle of memory that leads to faster access of their semantic representations. The present experiment was not designed to evaluate the relative explanatory power of the cross-language theories (i.e., the distributed feature and representation models) versus the more general proposals regarding semantic differentiation (e.g., dual-coding and context availability theories). However, if a general model can account for both the within and cross-language conceptual salience effect, it will not be necessary to postulate an additional cross-language mechanism like the one in the distributed feature model. This is not to say that the distributed feature and representation models should be discarded entirely, however, because the distributed models explain some effects that are not explained by the general models (e.g., word association and cognate effects).

Revised Hierarchical Model. Finally, the revised hierarchical model predicts that L1 to L2 translation is more likely than L2 to L1 translation to engage meaning. Therefore, there should be larger effects of conceptual salience (a semantic variable) in the L1 to L2 direction of translation than in the L2 to L1 direction.

Ambiguity Effects

In this study, “ambiguity” will be used to refer to a set of factors that encompasses both conceptual and lexical ambiguity. When these factors are referred to individually, they will be labeled number of meanings and number of forms for the meaning selected, respectively.

Tokowicz and Kroll Model. It was hypothesized that words would be translated more slowly as the ambiguity (number of meanings and number of forms for the meaning selected) across languages increased because the translation task in this experiment requires that a single node in memory be selected for production. Furthermore, translation from L1 to L2 should be influenced by both lexical and conceptual ambiguity whereas translation from L2 to L1 may be influenced only by lexical-level ambiguity. The result would be ambiguity effects limited to multiple forms from L2 to L1 if this direction is lexically mediated. The model predicts that the effects of ambiguity will be greater when there is ambiguity at multiple levels in the system. That is, words with both multiple meanings and multiple translations should be translated more slowly than words with ambiguity at only one level. However, there are too few words that have ambiguity at both levels to examine this prediction. In addition, the translation of words with multiple forms will be slower than that of words with multiple meanings because the model proposes that conceptual ambiguity can be overcome by a dominant translation, but that the reverse is not possible.

Interaction between Conceptual Salience and Ambiguity

Tokowicz and Kroll Model. The model proposed by Tokowicz and Kroll predicts that the effects of conceptual salience will only be evident when ambiguity is present. The model proposes that it is the slowing of the lexical selection process due to ambiguity that allows the effects of conceptual salience to become evident. This interaction is not predicted by any other model.

Cognate Effects

Distributed Feature Model. Another factor that has been shown to influence translation performance is cognate status (e.g., De Groot, 1992; Sánchez-Casas et al., 1992). In general, cognates should be translated more quickly than noncognates. This is because of their more complete feature overlap across languages (e.g., De Groot, 1992). According to the distributed feature model, cognates share all of their conceptual features across languages and therefore are translated more quickly than noncognates. Thus, the distributed feature model focuses on the conceptual level of representation. Similarly, De Groot and Nas (1991) concluded that bilinguals transfer their L1 meaning to L2 for cognate words and therefore cognates share a concept in the two languages. This conclusion was based on the finding that only cognates showed cross-language semantic priming when the primes were masked.

Distributed Representation Model. An alternative explanation of cognate effects is possible, however. Because cognates share both lexical form and meaning across languages, their overlap at the form level may be at least partly responsible for the effects. In particular, the word form itself may serve as a memory cue that results in the faster access of the translation. Indeed, cued translation, a task in which the first letter of the translation is given, leads to faster reaction times (De Groot, 1992). Thus, the cognate effects may come from the form level as well as the meaning level. The distributed representation model acknowledges the more complete lexical overlap for cognates than noncognates. Therefore, the focus of this model is on the overall similarity between the stimulus and the response, in terms of lexical features as well as conceptual features.

Although it is not possible to use the data of the present study to determine which model is more appropriate, it is most likely the case that the model that takes both levels of representation into account will be able to explain more of the results. This is because there are cases in which there is overlap at only one level and this latter model allows for such distinctions.

Interaction between Conceptual Salience, Ambiguity, and Cognate Status

Consistent with the idea that cognates are processed differently than noncognates (perhaps due to their differential meaning overlap as hypothesized by De Groot, 1992 and Van Hell & De Groot, 1998b), a post-hoc analysis of cognate effects by Tokowicz and Kroll showed that the concreteness effect was only observed for cognates when they had one translation, the reverse of what was found for noncognates. These data suggest that the interaction of concreteness and ambiguity is conditional on cognate status. This possibility will be evaluated in this experiment.

Method

Participants

The participants in this experiment were 26 Dutch-English bilinguals who were students at the University of Nijmegen, Nijmegen, The Netherlands. They were paid or given extra credit towards their studies in psychology for their participation. Their language history questionnaire data are shown in Table 5.

Stimuli

The stimuli were 548 English words and their translations in Dutch. One hundred and twenty-four items served as fillers; norms had been collected previously on the remaining 440 translation pairs that had been used in several past studies (De Groot, 1992; De Groot et al., 1994; Van Hell, 1998a, 1998b). These norms included printed word frequency (in Dutch and English; the CELEX database; Baayen, Piepenbrock, & Gulikers, 1995) and word concreteness and context availability in Dutch and English (De Groot, 1992). The Dutch word frequencies were obtained from 930 fiction and non-fiction books (approximately 30% fiction, 70% non-fiction) published between 1970 and 1988. The English word frequencies were obtained from 284 written texts taken from the COBUILD/Birmingham corpus

Table 5

Language History Questionnaire Data from the Participants of Experiment 1

Measure	Results
Age (years)	21.6 (2.1)
Age Began L2 (years)	9.9 (3.2)
Time Studied L2 (years)	9.7 (3.7)
L2 Immersion Experience (months)	2.3 (4.3)
L2 Reading Ability	8.4 (.9)
L2 Writing Ability	6.9 (1.7)
L2 Conversation Ability	7.4 (1.6)
L2 Speech Comprehension Ability	8.8 (1.1)

Note. Standard deviations are given in parentheses. Reading, writing, conversational, and speech comprehension ability were rated on a 10-point scale where 1 indicated the lowest level of ability and 10 indicated the highest level of ability.

(Sinclair, 1987). Because the frequency distributions were non-normal, the logarithmic transformations of the frequencies were used in the regression analyses.

See the General Method section for information about the number of translations and cognate norms that were also collected. The properties of the critical stimulus words are given in Table 6.

Table 6

Properties of the Critical Stimuli from Experiment 1 by Language of Stimulus

	Language of Stimulus	
Measure	Dutch	English
Length (number of letters)	5.8 (2.0)	5.7 (1.8)
Printed Word Frequency (per million)	107.3 (266.0)	132.5 (266.7)
Concreteness Rating	4.8 (1.7)	5.1 (1.6)
Context Availability Rating	4.7 (.9)	5.0 (.9)

Note. Standard deviations are given in parentheses. The ratings were performed on a 7-point scale on which seven indicated the highest level of concreteness/context availability.

Procedure

All participants were tested individually and were given verbal instructions by the experimenter, followed by additional instructions on the computer screen. Participants translated aloud visually-presented words from L1 to L2 and L2 to L1. The order of the directions of translation was counterbalanced across participants. Practice trials were given prior to the critical trials in each direction of translation. Prior to the presentation of each stimulus, a fixation point was presented at the center of the computer screen until the participant initiated the beginning of the trial by pressing a key on the button box. The stimuli replaced the fixation and remained on the computer screen until the participant made a verbal response. Participants were instructed to respond as quickly and accurately as possible, and to indicate when unaware of the correct translation of the presented word by saying “no” or its equivalent in Dutch. Their verbal responses were tape recorded then later coded for accuracy. Reaction time in milliseconds (ms) was recorded by the computer voicekey from the onset of stimulus presentation to the onset of the verbal response. The PsyScope computer program was used to present the stimuli in a random order and to record reaction time (Cohen, MacWhinney, Flatt, & Provost, 1993).

Data Trimming

Data from two participants were lost due to equipment failures; the final analyses included data from 24 participants. All participants translated accurately at least 60 % of the time in both directions of translation (including filler trials). Only correct responses on critical trials were included in the reaction time analyses; the accuracy analyses included data from all critical trials. Accuracy was coded using the norms established in the number of translations task; only those words considered accurate translations during the norming task were considered correct.

Reaction times that were faster than 300 ms or slower than 3000 ms were removed from the analyses, resulting in the removal of 4.5 and 3.6 % of the data in the L1 to L2 and L2 to L1 directions of translation, respectively. In addition, reaction times that were 2.5 SDs above or below the participant’s mean RT were excluded from the analyses and were treated as missing values.[13] These criteria were applied to the two directions of translation separately and resulted in the removal of 3.3 and 3.1% of the data for L1 to L2 and L2 to L1 translation, respectively. Finally, the voice key failed on .3 and .1 % of the trials in the two directions of translation. These data were treated as missing values.

Results and Discussion

There were several questions of interest in the present study. Separate hierarchical regression analyses were used to analyze the data relevant to each question. Furthermore, direction of translation was included as a repeated measure to determine whether the influence of the primary factor on translation differed in the two directions. In each analysis, the length and the logarithmic transformation of the frequency for both the stimulus and the response served as covariates. All factors previously shown to affect translation performance were entered on the first step of the analysis along with the covariates; the critical variable or set of variables for each analysis was always entered on the final step of the analysis. Because participants' responses were idiosyncratic, the regression analyses were run for each participant individually, then the average change in R^2 for the final step of the analyses across participants was submitted to a repeated measures ANOVA in which direction of translation was the repeated measure. In the analyses in which direction of translation was

[13] Deviant responses were removed because the theories presented here are aimed at predicting typical performance. Extremely fast scores are thought to reflect anticipatory processes and extremely slow scores are attributed to lapses of attention, and therefore do not reflect the processes of interest (Ratcliff, 1993).

the critical factor, a one-group *t*-test with a population mean of zero was performed instead of an ANOVA.

In order to demonstrate the nature of the effects, for each analysis, estimated reaction times were calculated using the regression equations. High (*M* + 1 *SD*) and low (*M* - 1 *SD*) values of the critical variable(s) for each analysis were entered, following the procedure suggested by Aiken and West (1991); the average value was entered for all other variables. High and low values were calculated using the same stimuli that were entered into the regression analyses. These means were calculated for each participant individually, then the average across participants was calculated. This average plus and minus one standard deviation from the mean were the high and low values that were subsequently entered into the regression equation. When there was a theoretical reason to use other values, as for the number of meanings or forms which can only assume integer values, the more appropriate values were entered.

Preliminary Analyses

Correlations. The correlations between the predictor variables were calculated for each participant including correct responses only. Although intercorrelations can be calculated based on the stimulus set, some intercorrelations are specific to the response the participant made (e.g., the cognate rating of the translation pair). Therefore, the intercorrelations were determined on a participant-by-participant basis. Because correlation distributions are non-normal, the correlations were transformed using Fisher's z' transformation (Cohen & Cohen, 1983, p. 53). The average z's across participants were then calculated and tested against a population mean of zero. Table 7 gives the average intercorrelations across participants that were obtained by converting the average z's to correlation values.

Concreteness and context availability were highly correlated in this stimulus set, as has been reported in past research (e.g., Schwanenflugel et al., 1988, Experiments 2 and 3; Schwanenflugel & Shoben, 1983, as cited above). As a result, it would not be possible to evaluate the effects of these two factors independently because the two predictors would account for largely the same portion of the variance. Therefore, both concreteness and context availability were included as a set, so that the joint contribution of the set of predictors could be evaluated. As mentioned above, the set of factors will be referred to as conceptual salience.

Similarly, the number of meanings and total number of translations a word had were highly correlated because words with more meanings tended to have more translations overall. In order to determine whether there are differential effects of multiple meanings versus multiple translations, the number of meanings and the number of forms (translations) for the meaning of the word that was selected for translation, were treated as a set of predictors known as “ambiguity”. Because these two measures were not correlated, it is possible to examine their independent contributions, even when analyzed as a set of factors. Consider, for example, the English word “meeting” which has two interpretations. The encounter meaning has one translation in Dutch, “ontmoeting”. In contrast, the assembly meaning has two translations in Dutch, “bijeenkomst” and “vergadering”. Therefore, when the encounter meaning was selected, the word was considered a two meaning, one translation

Table 7

Average Intercorrelations Between Factors from Dutch to English and English to Dutch Translation

Factor	1	2	3	4	5	6
			Dutch to English Translation			
1. Concreteness Rating	--	.82**	.14	-.07	-.16	-.14
2. Context Availability Rating		--	.10	-.12	-.17	-.12
3. Cognate Rating			--	-.05	-.11	-.10
4. Number of Meanings				--	.59**	-.05
5. Number of Translations					--	.78**
6. Forms for Meaning Selected						--
			English to Dutch Translation			
1. Concreteness Rating	--	.80**	.11	-.16	-.26	-.20
2. Context Availability Rating		--	.09	-.28	-.27	-.16
3. Cognate Rating			--	-.00	-.12	-.13
4. Number of Meanings				--	.55**	.01
5. Number of Translations					--	.83**
6. Forms for Meaning Selected						--

Note. ** $p < .01$

word. When the assembly meaning was selected, it was considered a two meaning, two translation word.

Because number of meanings and the number of forms for the meaning selected are not correlated, the inclusion of both factors did not result in problems interpreting the independent contribution of the predictors. Furthermore, it is assumed that in cases in which words have multiple translations, the translation is not selected until the meaning has been chosen (Tokowicz & Kroll, submitted).

Although most words had only one meaning and one translation, there were some words with multiple forms and one meaning or multiple meanings and one form for that meaning. Table 8 shows the average number of words in each category that were translated correctly by the first three participants. This average is representative of the larger group of

Table 8

The Average Number of Words Translated Correctly by the First Three Participants by Number of Meanings and Number of Forms for the Meaning Selected

	Number of Forms for the Meaning Selected			
Number of Meanings	1	2	3	4
1	285	41	6	2
2	21	4	--	--
3	2	--	--	--

participants. As can be seen, there are very few words with more than two forms or meanings, and only four words on average with both multiple forms and meanings. Therefore, the three most common categories are one meaning/one form, one meaning/two forms, and two meanings/one form. For the analyses in which ambiguity varied, these values were entered into the regression equation to determine estimated values. These three categories of words were then compared.

Proportion of Variance Accounted for by Each Predictor. In a preliminary set of analyses, the proportion of variance accounted for by each predictor or set of predictors was determined using the mean change in R^2 across participants. In each analysis, the length of the stimulus, the length of the response, and the logarithmic transformation of the stimulus and response frequencies were included as covariates. The analysis for cognate rating included no other predictors. The analyses for direction of translation, concreteness, and ambiguity included cognate rating as an additional covariate. The results of these analyses are given in Table 9 where the predictors are presented in the order of importance in terms of the proportion of variance accounted for.

Scatterplots. The standardized predicted values of the dependent variable Y and the residuals (i.e., the difference between the actual and predicted values of Y) from the above analyses were saved and then plotted as scatterplots to verify the validity of using linear regression analyses, following the procedure suggested in Cohen and Cohen (1983, p.126). For purposes of brevity, the scatterplots from the first three participants are presented in the appendix (see Appendix D). Because these scatterplots are representative of those from all participants, the additional plots were omitted. These plots suggested that there were no

Table 9

Average Change in R^2 Across Participants Due to the Inclusion of Each Predictor Variable

Predictor	Average R^2 Change
Direction of Translation	.026 ± .007
Dutch to English Translation	
Cognate Rating of Translation Pair	.050 ± .011
Conceptual Salience of Stimulus	.032 ± .024
Ambiguity	.024 ± .005
English to Dutch Translation	
Cognate Rating of Translation Pair	.067 ± .008
Conceptual Salience of Stimulus	.044 ± .042
Ambiguity	.029 ± .006

Note. Values indicate the mean plus or minus the standard error of the mean. Analyses separating the influence of concreteness and context availability were also run. These analyses showed that context availability was a slightly better predictor of translation latency; the average changes in R^2 for context availability and concreteness, respectively, were .027 ± .005 versus .020 ± .005 for L1 to L2 translation and .035 ± .009 versus .018 ± .005 for L2 to L1 translation.

problems of heteroscedasticity or curvilinearity; linear regression analyses were therefore used to analyze these data.

Effects of Conceptual Salience

The first question of interest in this experiment was whether conceptual salience would account for additional variance after other factors that have been shown to influence bilingual language processing (e.g., length, frequency, cognate status, and ambiguity) were accounted for (see Table 10 for an overview of this analysis).

Table 10

Overview of the Hierarchical Regression Analysis Testing for Conceptual Salience Effects

Step	Variable
1	Stimulus length (number of letters)
	Response length (number of letters)
	Log frequency of stimulus
	Log frequency of response
	Number of meanings
	Forms for meaning selected
	Cognate rating of translation pair
	Number of meanings x cognate rating
	Forms for meaning selected x cognate rating
2	Concreteness rating of stimulus
	Context availability rating of stimulus

The distributed models and general models of conceptual salience predict that conceptual salience effects should always be present. In contrast, the revised hierarchical model predicts that these effects should be present only for L1 to L2 translation.

This analysis showed that the effect of conceptual salience was reliable after ambiguity was taken into account [mean $\Delta R^2 = .029$, $F(1, 23) = 74.2$, $p < .01$] and there was no difference between the two directions of translation. In particular, the higher the conceptual salience, the faster the response. High (*M* + 1 *SD*) and low (*M* - 1 *SD*) values of concreteness/context availability were entered into the regression equation yielding estimated RTs at high and low levels of conceptual salience. Due to the confounding of concreteness and context availability, it was assumed that high levels of concreteness corresponded to high levels of context availability; analyses including high levels of one factor and low levels of the other were not performed. The estimated values were 902 ms for words high in conceptual salience (e.g., "farm") and 977 ms for words low in conceptual salience (e.g., "curse"). Thus, the results of this analysis show that there is an overall effect of conceptual salience that remains even after the effects of ambiguity have been taken into account.

Furthermore, conceptual salience influenced translation accuracy, [mean $\Delta R^2 = .039$, $F(1, 23) = 68.1$, $p < .01$]; there were no differences between the two directions of translation. Words higher in conceptual salience were translated more accurately than those low in conceptual salience (79 versus 48 %). These findings are consistent with the predictions of both the distributed feature model and of the more general models (e.g., dual coding and context availability theories) that propose that the meanings of words high in conceptual salience are more readily accessed than the meanings of words low in conceptual salience.

The finding that conceptual salience predicts translation performance is consistent with the predictions of the distributed feature and representation models as well as the dual-coding and context availability theories. Interestingly, there were effects of conceptual salience in both directions of translation. This finding can be taken as evidence against the revised hierarchical model which suggests that these effects should only occur (or should be greater) during L1 to L2 translation. However, this result is consistent with models that do not propose a difference between the two directions of translation in the extent to which they involve meaning access (e.g., the distributed feature model, De Groot, 1992; the distributed representation model, Van Hell & De Groot, 1998a). Finally, the finding of an overall effect of conceptual salience is similar to the results reported by Tokowicz and Kroll who found a main effect of concreteness such that concrete words were translated more quickly and accurately than abstract words.

Ambiguity Effects

The second question of interest in this experiment was whether ambiguity accounts for additional variance once conceptual salience has been taken into account (see Table 11 for an overview of the analysis). The only model that addresses the issue of cross-language ambiguity is the one presented by Tokowicz and Kroll. This model predicts that there should be an overall effect of ambiguity. However, this effect may be different for the two directions of translation; L1 to L2 translation should be influenced by both number of meanings and number of forms for the meaning selected, whereas L2 to L1 translation may be influenced only by the number of forms for the meaning selected because it is not necessary for

Table 11

Overview of the Hierarchical Regression Analysis Testing for Ambiguity Effects

Step	Variable
1	Stimulus length (number of letters)
	Response length (number of letters)
	Log frequency of stimulus
	Log frequency of response
	Concreteness rating of stimulus
	Context availability rating of stimulus
	Cognate rating of translation pair
	Concreteness x cognate rating
	Context availability x cognate rating
2	Number of meanings
	Forms for meaning selected

conceptual mediation to occur during this direction of translation. The result of this difference would be a smaller ambiguity effect for L2 to L1 translation.

The analysis showed that the effect of ambiguity was reliable [mean $\Delta R^2 = .025$, F (1, 23) = 37.1, $p < .01$] and that there was no difference between the two directions of translation. High and low values of the number of meanings and number of forms for meaning selected were entered into the regression equation to determine estimated values at different levels of the predictors (see Table 12). Because the estimated values using the mean number of meanings and forms plus and minus one standard deviation were not meaningful, one was used as the low value and two was used as the high value. An example of a stimulus

Table 12

Estimated Reaction Times (ms) and Accuracy Values (%) as a Function of Number of Meanings and Number of Forms for the Meaning Selected

	Forms for Meaning Selected	
Number of Meanings	One	Two
One	925 (64.9)	988 (56.7)
Two	973 (62.8)	--

with few meanings and few forms is the word "dress" which translates to "jurk" in Dutch. A word with one meaning but multiple forms is the word "town" which translates to both "stad" and "plats" in Dutch. Finally, an example of a stimulus with multiple meanings but one form for the meaning selected is the word "meeting" when its encounter interpretation is translated (and not its assembly interpretation) as "ontmoeting". Because there were very few words with multiple meanings and multiple translations, the estimated translation latency was not calculated for such words.

The data show that the greater the ambiguity, the longer the reaction time. This effect did not differ between the two directions of translation. The Tokowicz and Kroll model left the possibility open for there to be a difference between the two directions of translation in terms of the level of the ambiguity (i.e., lexical or conceptual). To test that possibility, a comparison of the additional amount of variance accounted for using the two ambiguity measures alone to predict performance in the two directions of translation was needed.

Therefore, two sets of hierarchical regression analyses were run. In the first analysis, the number of meanings were entered on the first level and the number of forms were entered on the second level. The reverse was done in the second analysis. Then the change in R^2 was evaluated. This measure allows the determination of whether adding the one variable alone accounts for additional variance. Both analyses were significant overall and there were no differences between the two directions of translation [mean $\Delta R^2 = .006$, $F(1, 23) = 19.54$, $p < .01$ for number of meanings, and mean $\Delta R^2 = .015$, $F(1, 23) = 20.84$, $p < .01$ for number of forms for meaning selected]. Therefore, there is no support for the possibility that the two directions differ in their sensitivity to different sources of ambiguity.

Furthermore, an inspection of the regression equation from the ambiguity analysis shows that both number of meanings and number of forms for the meaning selected have similar effects on translation performance. In both cases, RT increases as the number of meanings/translations increases. Furthermore, the procedure presented in Cohen & Cohen for comparing regression coefficients from the same regression equation (1983, pp. 479-500) was followed to compare the coefficients from the two factors. This procedure showed that there was not a significant difference in the size of the regression coefficient for the number of meanings and the number of forms for the meaning selected [mean *B*s = 47.994 and 63.182, respectively; $t(12) = -.69$, $p > .10$]. This suggests that the two factors have similar effects on translation performance. This finding does not support the prediction made by the Tokowicz and Kroll model that multiple forms for the meaning selected would slow performance more than multiple meanings.

There was also a reliable effect of ambiguity on accuracy [mean $\Delta R^2 = .012$, $F(1, 23) = 53.4$, $p < .01$], such that the greater the number of translations for a word, the less accurately it was translated (see Table 12 above). There were no differences between the two directions of translation. These findings are important because they suggest that ambiguous words are more difficult to translate even when all possible answers are accepted as correct. Presumably, this is due to the increased competition that is created when a word has multiple translations. These results are consistent with those of Tokowicz and Kroll who reported a main effect of ambiguity on translation latency and accuracy in the same direction as those reported here.

Interaction between Conceptual Salience and Ambiguity

Tokowicz and Kroll reported an interaction between concreteness and ambiguity such that concrete words were unaffected by ambiguity. In particular, there was no concreteness effect for unambiguous words, and a standard concreteness effect for ambiguous words. The product of ambiguity and conceptual salience was used to test for an analogous interaction in this experiment (see Table 13 for an overview of this analysis). This analysis showed that the interaction between conceptual salience and ambiguity was reliable [mean $\Delta R^2 = .023$, $F(1, 23) = 109.3$, $p < .01$] and there was no difference between the two directions of translation. The estimated translation latencies were calculated by entering low and high values (1 and 2, respectively) of number of meanings and number of forms for meaning selected, as well as conceptual salience (see Figure 4).

The estimated reaction times show that unambiguous words (words with one meaning and one form) are translated more quickly than either type of ambiguous word (multiple

Table 13

Overview of the Hierarchical Regression Analysis Testing for an Interaction Between Conceptual Salience and Ambiguity

Step	Variable
1	Stimulus length (number of letters)
	Response length (number of letters)
	Log frequency of stimulus
	Log frequency of response
	Concreteness rating of stimulus
	Context availability of stimulus
	Cognate rating of translation pair
	Concreteness x cognate rating
	Context availability x cognate rating
	Number of meanings
	Forms for meaning selected
	Number of meanings x cognate rating
	Forms for meaning selected x cognate rating
2	Number of meanings x concreteness
	Forms for meaning selected x concreteness
	Number of meanings x context availability
	Forms for meaning selected x context availability

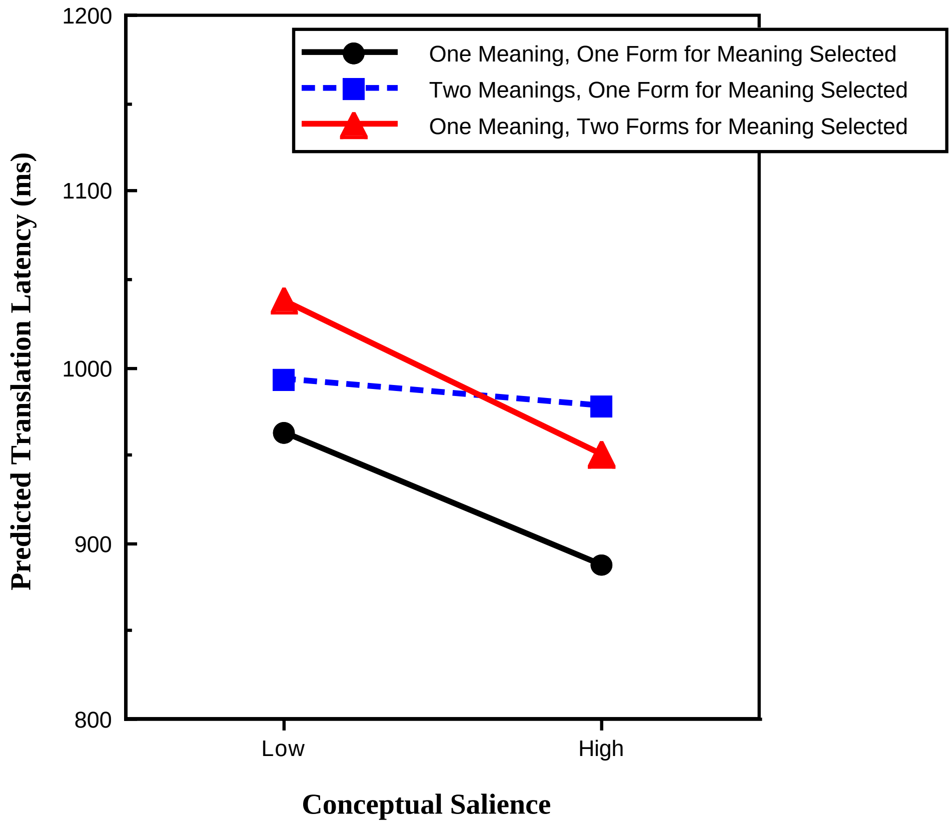


Figure 4. Estimated translation latencies (ms) as a function of conceptual salience, number of meanings, and number of forms for the meaning selected.

forms or multiple meanings). Furthermore, words with two forms were translated more slowly than unambiguous words, but this does not interact with conceptual salience. Words with two meanings were translated more slowly than unambiguous words but this difference was greater for words high in conceptual salience. These results are quite different from those observed by Tokowicz and Kroll in which the ambiguity effect was greater for words lower in concreteness, however cognates were not included as critical items in their study. Therefore, it is possible that the effect would be different for cognates and noncognates, as was suggested by the post-hoc analysis performed by Tokowicz and Kroll on the cognate items. The implications of this difference between the experiments will be addressed further in the section on the interaction between conceptual salience, ambiguity, and cognate status.

There was also a significant effect of the interaction between conceptual salience and ambiguity on accuracy [mean $\Delta R^2 = .020$, F (1, 23) = 73.2, $p < .01$]. Again, high and low values of the critical predictors were used to calculate estimated accuracy scores for each type of word (see Table 14). Words high in conceptual salience were translated more accurately than words low in conceptual salience. Words with multiple forms were translated less accurately than those with one form. Words with multiple meanings were generally translated less accurately than those with one meaning, however this difference was greater for words high in conceptual salience. Further, multiple forms made performance more error-prone than multiple meanings. The finding that meaning effects were greater for words high in conceptual salience may be due to the relative strength of the conceptual representations for these words; stronger concepts may provide more competition for other concepts. Further, the finding that multiple forms influenced words high and low in conceptual salience

Table 14

Estimated Accuracy Scores (%) as a Function of Conceptual Salience, Number of Meanings, and Number of Forms for the Meaning Selected

	Forms for Meaning Selected	
Number of Meanings	One	Two
	Low Conceptual Salience	
One	49.2	41.8
Two	47.7	--
	High Conceptual Salience	
One	80.8	66.2
Two	73.5	--

similarly is consistent with the idea that these words do not differ in the strength of their connections to their translations at the form level.

Cognate Effects

Past research has shown that cognates are translated more quickly than noncognates (e.g., De Groot, 1992; Sánchez-Casas et al., 1992). This analysis was conducted to determine whether the cognate ratings of a translation pair affected translation performance as predicted by the distributed feature and representation models (see Table 15 for an overview of the analysis).

Table 15

Overview of the Hierarchical Regression Analysis Testing for Cognate Effects

Step	Variable
1	Stimulus length (number of letters)
	Response length (number of letters)
	Log frequency of stimulus
	Log frequency of response
	Number of meanings
	Forms for meaning selected
	Concreteness rating of stimulus
	Context availability of stimulus
	Number of meanings x concreteness
	Forms for meaning selected x concreteness
	Number of meanings x context availability
	Forms for meaning selected x context availability
2	Cognate rating of translation pair

This analysis showed that the effect of cognate rating was reliable [mean $\Delta R^2 = .056$, F (1, 23) = 67.3, $p < .01$]; there were no differences between the two directions of translation. In particular, translation pairs that were higher in their rated cognate status (e.g., "blame"-"blaam") were translated more quickly than those lower in rated cognate status (e.g., "sky"-"lucht"; estimated values: 873 versus 1008 ms). There was also a significant effect of cognate rating on accuracy [mean $\Delta R^2 = .019$, F (1, 23) = 54.3, $p < .01$]; translation pairs higher in rated cognate status were translated more accurately than those lower in rated cognate status (73 versus 53 %).

The results of this analysis are consistent with the predictions of the distributed feature and representation models. However, these two models have somewhat different explanations of the cognate effect. Because the distributed feature model focuses on overlap at the conceptual level, it relies on the assumption that cognate words share more conceptual features than noncognates. The distributed representation model takes both conceptual and lexical-level overlap into account and therefore explains cognate effects on the basis of more similar meaning and form for these words across languages.

Interaction between Conceptual Salience, Ambiguity, and Cognate Status

The form of the interaction between concreteness and ambiguity observed in this experiment is different from the findings of Tokowicz and Kroll and does not support the predictions made by the model they presented. However, cognates were not included as critical items in that study. In the present study, translation pairs varied widely in their cognate status. Furthermore, as De Groot (1992) argued, there may be reason to believe that cognates are processed differently than noncognates. The final analysis was aimed at determining whether cognate status moderates the interaction between conceptual salience and number of translations (see Table 16 for an overview of this analysis). This possibility was suggested by the post-hoc analysis of cognates in the Tokowicz and Kroll study. That analysis showed that cognates and noncognates were differentially influenced by concreteness and ambiguity. For cognates, there was a larger concreteness effect for words with one translation whereas for noncognates there was a larger concreteness effect for words with more than one translation.

Table 16

Overview of the Hierarchical Regression Analysis Testing for an Interaction between Cognate Status, Conceptual Salience, and Number of Translations

Step	Variable
1	Stimulus length (number of letters)
	Response length (number of letters)
	Log frequency of stimulus
	Log frequency of response
	Concreteness rating of stimulus
	Context availability of stimulus
	Cognate rating of translation pair
	Concreteness x cognate rating
	Context availability x cognate rating
	Number of meanings
	Forms for meaning selected
	Number of meanings x concreteness
	Forms for meaning selected x concreteness
	Number of meanings x context availability
	Forms for meaning selected x context availability
	Number of meanings x cognate rating
	Forms for meaning selected x cognate rating
2	Cognate status x concreteness x number of meanings
	Cognate status x concreteness x forms for meaning selected
	Cognate status x context availability x number of meanings
	Cognate status x context availability x forms for meaning selected

This analysis showed a reliable interaction between conceptual salience, ambiguity, and cognate rating [mean $\Delta R^2 = .027$, F (1, 23) = 67.9, $p < .01$]; there was no difference between the two directions of translation. The probative procedure presented in Aiken and West (1991) was used to describe the nature of this interaction. Simple regression equations were generated from the combinations of high and low levels of conceptual salience ($M \pm 1$ SD) and ambiguity (1 and 2) across participants. Low and high values of cognate rating ($M \pm$ 1 SD) were then entered into the regression equations. The resulting graph (see Figure 5) shows the estimated reaction times at various levels of the independent variables.

Noncognates. The results show that for noncognates (translation pairs low in their cognate rating), the conceptual salience effect is present such that words high in conceptual salience were translated more quickly than words low in conceptual salience. Furthermore, this effect is similar for unambiguous words (few meanings, few forms) and words with multiple forms. However, there was no conceptual salience effect for words with multiple meanings. The results for unambiguous words are not consistent with those presented by Tokowicz and Kroll for words with one translation. Potential explanations for this discrepancy will be considered below.

The effect of ambiguity was present for words with multiple forms such that these words were translated more slowly than unambiguous words. However, the effect for words with multiple meanings interacted with conceptual salience, such that there was no effect of number of meanings for words low in conceptual salience, whereas there was a large effect of number of meanings for words high in conceptual salience, such that words with multiple meanings were translated more slowly than unambiguous words. Interestingly, these words with multiple meanings were also translated more slowly than words with multiple forms.

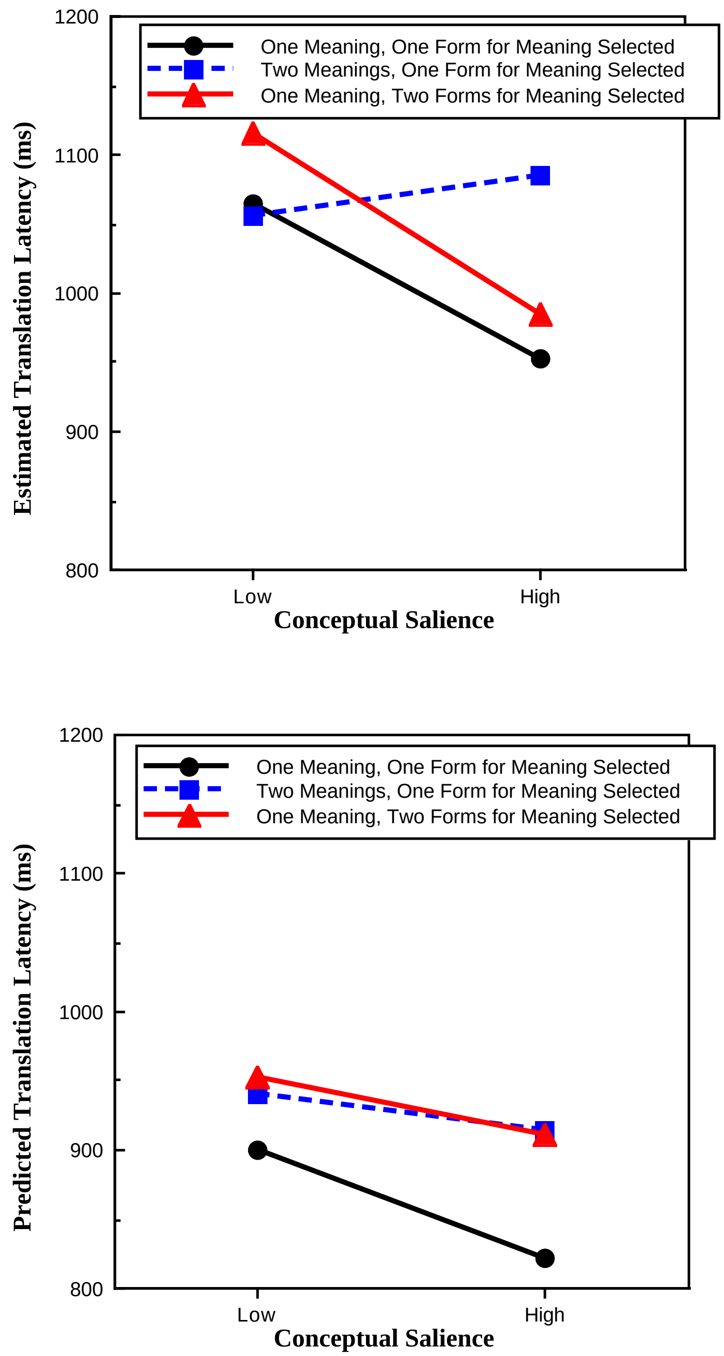


Figure 5. Estimated translation latencies (ms) at high and low levels of conceptual salience, number of translations, and cognate ratings The top panel shows translations low in cognate ratings (noncognates); the bottom panel shows translations high in cognate ratings (cognates).

These findings are important because they show that ambiguity can be examined with respect to the type of ambiguity (i.e., multiple forms or meanings), and furthermore that the type of ambiguity has implications for language processing. Interestingly, the number of meanings may also be relevant for performance within a language as well (see Experiment 2).

Cognates. First, there was an overall effect of cognate status such that cognates were translated more quickly than noncognates for all conditions. For cognate translations, a quite different pattern of results emerged. First, the effects tended to be smaller in general than for the noncognates. The effect of conceptual salience was small for the unambiguous words and negligible for the ambiguous words.

The ambiguity effect was similar for words with multiple meanings and words with multiple forms, which contrasts with the results for the noncognates. This ambiguity effect was larger for words high in conceptual salience than those low in conceptual salience. This finding is similar to post-hoc results reported by Tokowicz and Kroll for cognates which showed a larger ambiguity effect for concrete than abstract words.

There was a similar interaction effect for translation accuracy [mean $\Delta R^2 = .018$, F (1, 23) = 52.0, $p < .01$]. See Table 17 for the estimated accuracy values. The results of this analysis show that noncognates high in conceptual salience with multiple forms are more difficult to translate than other noncognates high in conceptual salience. The same was true for cognates low in conceptual salience. Further, cognates high in conceptual salience with multiple meanings were more difficult to translate than other cognates high in conceptual salience. Finally, there appear to be no differences among noncognates low in conceptual salience.

Table 17

Estimated Accuracy Scores (%) as a Function of Cognate Rating, Conceptual Salience, Number of Meanings, and Number of Forms for the Meaning Selected

	Cognate Rating			
	Low		High	
	Forms for Meaning Selected			
Number of Meanings	Few	Many	Few	Many
	Low Conceptual Salience			
Few	30.5	29.1	70.6	53.3
Many	30.7	--	75.9	--
	High Conceptual Salience			
Few	72.6	58.6	85.4	87.2
Many	68.1	--	77.5	--

Direction of Translation Effects

Past research has shown that translation from L1 to L2 is often performed more slowly and less accurately than translation from L2 to L1 (e.g., Kroll & Stewart, 1994). This analysis was aimed at determining whether this translation asymmetry was obtained in this experiment as predicted by the revised hierarchical model (see Table 18 for an overview of the analysis).

Table 18

Overview of the Hierarchical Regression Analysis Testing for Direction of Translation Effects

Step	Variable
1	Stimulus length (number of letters)
	Response length (number of letters)
	Log frequency of stimulus
	Log frequency of response
	Cognate rating of translation pair
2	Direction of translation

The results of this analysis showed that translation latencies were influenced by the direction of translation, mean $\Delta R^2 = .026$, t (23) = 3.8, $p < .01$; however, translation was performed more *slowly* in the L2 to L1 direction than in the L1 to L2 direction (estimated latencies: 987 versus 926 ms), the opposite of what has been found in past research. The direction of translation also affected translation accuracy, mean $\Delta R^2 = .005$, t (23) = 4.6, $p < .01$, such that translation was more accurate in the L1 to L2 than the L2 to L1 direction of translation (56 versus 55 %). Again, this result is the reverse of what is typically observed (e.g., Kroll & Stewart, 1994).

These results are quite surprising given that past research has typically shown either an asymmetry such that translation from L1 to L2 takes longer to perform and is more error prone than translation from L2 to L1, or that there is no difference between the two directions of translation. Although these results are not consistent with the predictions of the revised hierarchical model, they are also not consistent with the predictions of the distributed models mentioned earlier. This is because the distributed models predict no difference in reaction time for the two directions of translation. Furthermore, the Tokowicz and Kroll model allowed for the possibility that translation from L2 to L1 would be performed more quickly than translation from L1 to L2, or that the two directions would be similar; these predictions were not supported.

What does the reversal of the translation asymmetry tell us about how translation is performed? Recent research suggests that language dominance may be more important for predicting performance on the translation task than the order in which languages are learned (e.g., Elsinger, 1996; Heredia, 1997). In these studies, it was shown that the first language can become the less dominant language if a bilingual spends sufficient time in the L2 environment or if the circumstances are sufficient to shift the bilingual's “language mode” so that the L2 is more active (e.g., Grosjean, 1985, 1997). This shift may also be attributable to the domain of the task that is being performed (e.g., Elsinger, 1996). Because the bilinguals in this study were highly fluent in L2 and were accustomed to using English in a university setting, it may be that this context coupled with the presence of an English-speaking experimenter and the knowledge that the task required the participants to speak English fluently, could have made English temporarily more dominant than Dutch. This explanation is quite unlikely, however, because the bilinguals in this population are dominant in Dutch according to their fluency ratings and they use Dutch more often than English in their daily lives.

A more likely explanation for the reversal of the translation asymmetry is that the order in which the directions of translation were performed influenced the direction of the asymmetry. Michael (1998, Experiment 2) tested a group of Dutch-English bilinguals from the same population as those who participated in the present experiment. In that study, there was also a reversal of the typical translation asymmetry, such that translation from L2 to L1 was performed more slowly than translation from L1 to L2. A post-hoc analysis showed that this reversal was due to those participants who performed L1 to L2 translation prior to L2 to L1 translation; participants who performed L2 to L1 translation first showed no translation asymmetry.

A post-hoc analysis using direction of translation order as a grouping factor showed that the bilinguals who performed L1 to L2 translation first showed a significant effect of direction of translation on RT, $sr = -.14$, $t\,(11) = -3.584$, $p < .01$, such that translation from L2 to L1 was performed more slowly than translation from L1 to L2. In contrast, bilinguals who performed L2 to L1 translation first showed no effect of direction of translation on reaction time, $sr = -.05$, $t\,(11) = -1.6$, $p > .10$.

A potential explanation for this order effects is that there is a criterion level that is set for the level of activation that is required for output. When translation from L1 to L2 is performed first, the criterion is set low because L2 has less activation than L1. The high activation of L1 is sufficient to meet this low criterion. However, when L2 to L1 translation is performed first, the criterion is set high because L1 has a relatively high level of activation. Then, this criterion is too high to be met by the L2 words in L1 to L2 translation, thereby lengthening reaction time and causing a reversal of the translation asymmetry.

A similar order effect was reported by Kroll and Merves (1986; see Figure 6) who found that the concreteness effect depended on the nature of the lexical decision task. When the type of word was blocked and concrete words were presented first (low criterion), there

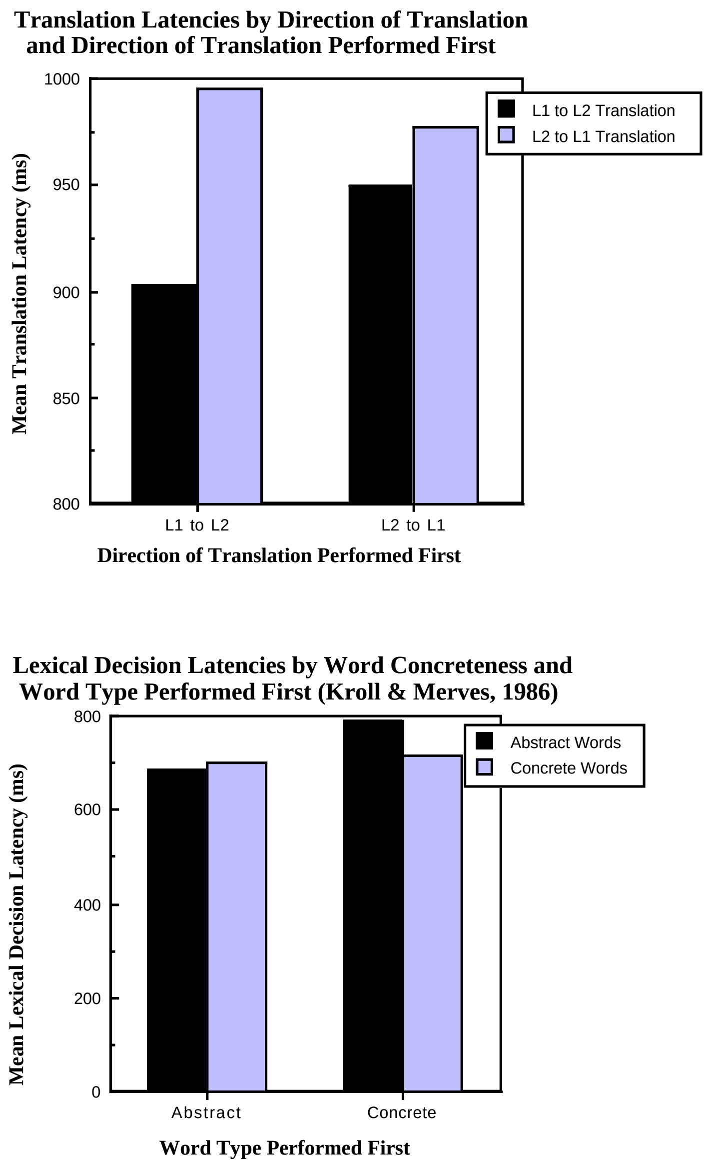


Figure 6. Translation latencies from a post-hoc analysis of Experiment 1 data as a function of the direction of translation performed first (top panel) and lexical decision latencies from Kroll and Merves (1986) as a function of word type performed first (bottom panel).

were strong concreteness effects, such that abstract words were responded to more slowly than concrete words. However, when abstract words were presented first (high criterion), there was no concreteness effect. The most important implication of these findings is that the context in which a task is performed can influence the pattern of results. When the pattern of results is used to determine the validity of models of language processing, it is critical to determine and explain which effects are due to the context and which effects are due to a property of the language processing system.

Conclusion

In sum, the results of this experiment show that multiple forms and multiple meanings have distinct effects on translation performance; although ambiguity generally slowed translation performance, the effects of multiple meanings were reduced for cognates and enhanced for concrete words, whereas the effects of multiple translations were relatively constant across word types. Thus, the effects of ambiguity are dependent on the conceptual salience and cognate rating of the translation pair. Tokowicz and Kroll reported that ambiguity affected translation performance and that concreteness effects were moderated by ambiguity. Although the same was true in the present study, the particular patterns of results are not the same. How can the results of that study be reconciled with those of the present experiment? The differences in the results may be due to the manner in which ambiguity was determined in the two studies. In the Tokowicz and Kroll study, directionality of ambiguity was not taken into account. However, there is one condition that is the same in the two experiments that did not show the expected pattern of results based on the prior study. Unambiguous words in this study showed an effect of conceptual salience whereas they did not in the Tokowicz and Kroll study. It is not clear why this is the case; however Tokowicz and Kroll dichotomized concreteness and this procedure may have led to an apparent null effect. Furthermore, although Tokowicz and Kroll matched words in the critical conditions as closely as possible in length and frequency, they did not use a covariance procedure to remove the effects of other factors that may have obscured the effect of concreteness.

The results of this experiment are not consistent with the predictions made by the Tokowicz and Kroll model. First, it was predicted that the difference between words low and high in conceptual salience would be greater for ambiguous words, and in particular for words with multiple forms. This was not true for either cognates or noncognates. It was also predicted that multiple meanings would not slow translation performance as much as multiple forms. However, this was only the case for noncognates low in conceptual salience.

The assumption of the model presented by Tokowicz and Kroll that only additional forms made the meaning resolution necessary to complete during translation is not consistent with the data from this experiment. The only effect of multiple forms is an overall addition of reaction time that is fairly constant across the conditions. In contrast, the effects of multiple meanings vary with the type of word being translated. The main effect that a model of these effects would have to explain is the finding that multiple meanings do not influence performance for low conceptual salience words, but they slow performance for high conceptual salience words. Perhaps this is because the words high in conceptual salience with multiple meanings will have two fairly active meanings that will compete more than multiple meanings for words low in conceptual salience. A final additional assumption would need to be made regarding the way that the translation task is performed. When a high conceptual salience word has multiple meanings, the relatively strong activation of a single lexical alternative for output does not resolve the ambiguity at the conceptual level. The main weakness of the Tokowicz and Kroll model in its original form is that it assumed that the activation at the lexical level could be sufficient to bypass the final resolution of the meaning, which appears not to be true.

A Revised Model of Conceptual Salience and Ambiguity Effects

The following are a new set of assumptions for a model like the one proposed by Tokowicz and Kroll (see Figures 7, 8, and 9). This revised model differs from the model proposed by Tokowicz and Kroll in that it is a more serial model in which a meaning is first selected and then the forms that correspond to that meaning are activated. These models illustrate the assumptions outlined in the text. The Dutch words at the lemma level are the translation equivalents of the English words shown at the lemma level. First, it is assumed that meaning selection is necessary to take place during translation. It is also assumed that word meanings are activated more quickly the higher they are in conceptual salience. When a high conceptual salience word has multiple meanings, these meanings will be relatively strong and therefore will compete with each other. This will lead to the slowing of the meaning resolution process. In contrast, when a low conceptual salience word has multiple meanings, assuming that one meaning has (slightly) more activation than the other, the secondary one may not serve as a competitor due to low activation. That is, there is a certain criterion a potential competitor must meet in order to compete for activation. Therefore,

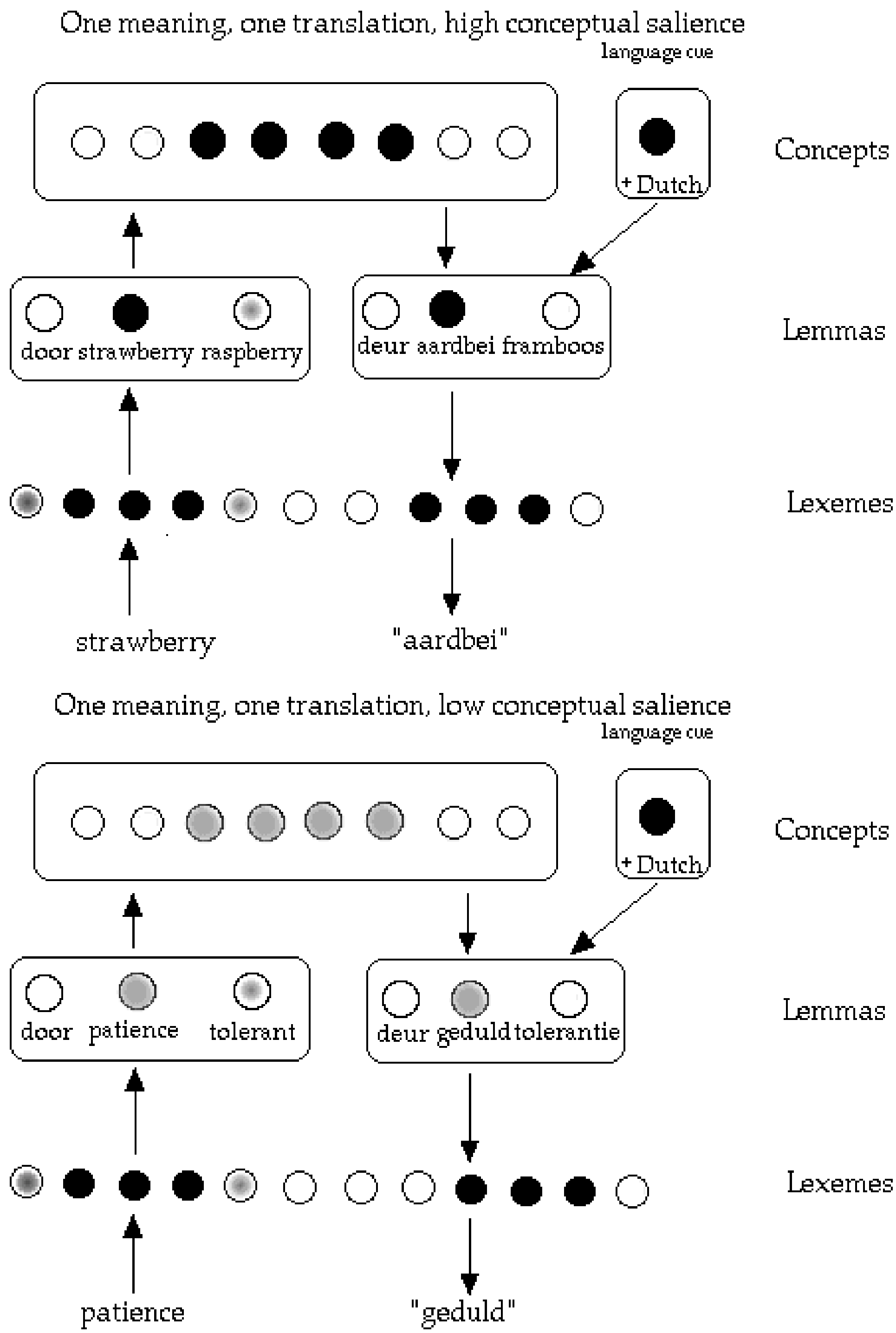


Figure 7. Revised model of conceptual salience and ambiguity effects for words with one meaning and one translation. The top panel shows words high in conceptual salience and the bottom panel shows words low in conceptual salience.

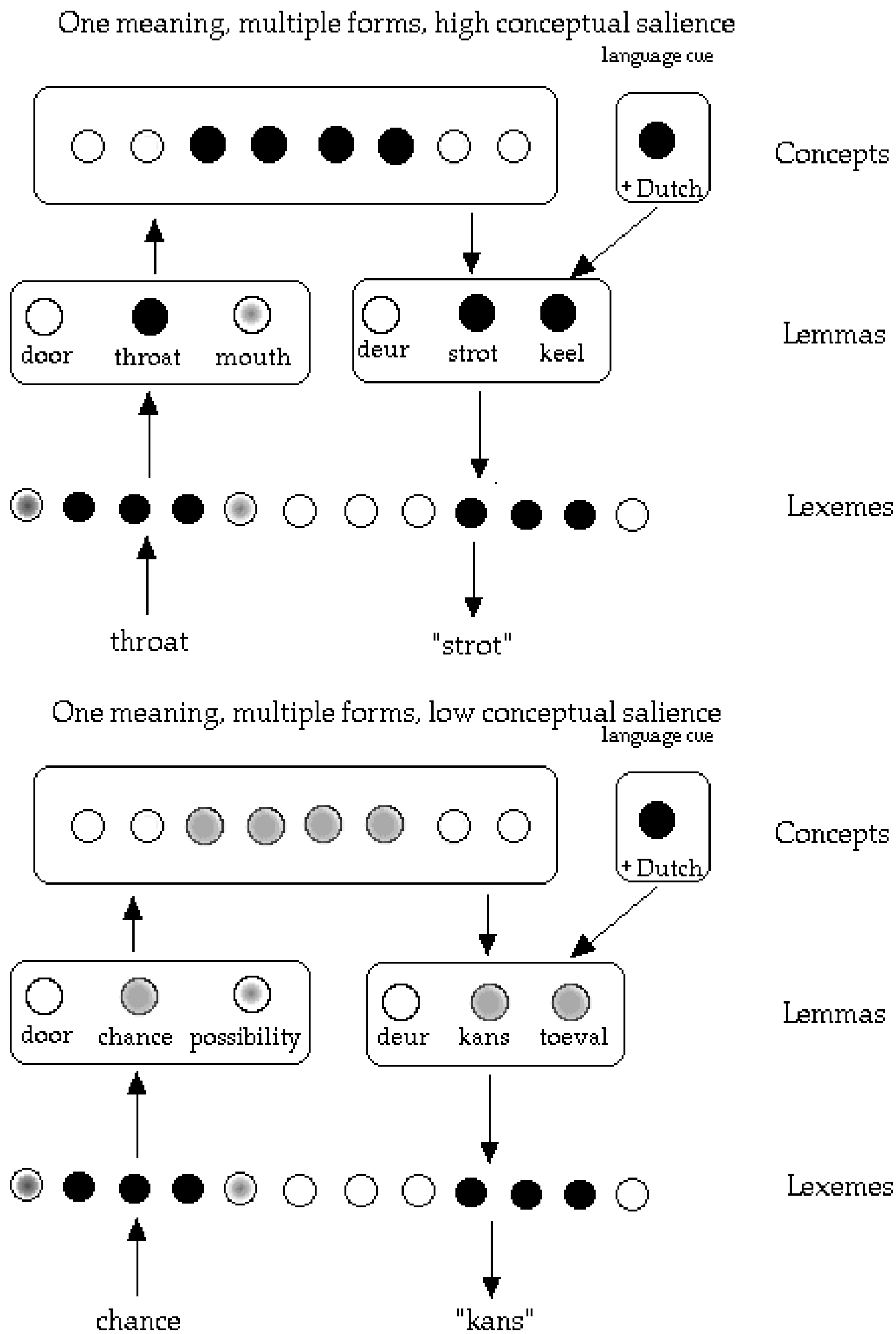


Figure 8. Revised model of conceptual salience and ambiguity effects for words with one meaning and multiple forms. The top panel shows words high in conceptual salience and the bottom panel shows words low in conceptual salience.

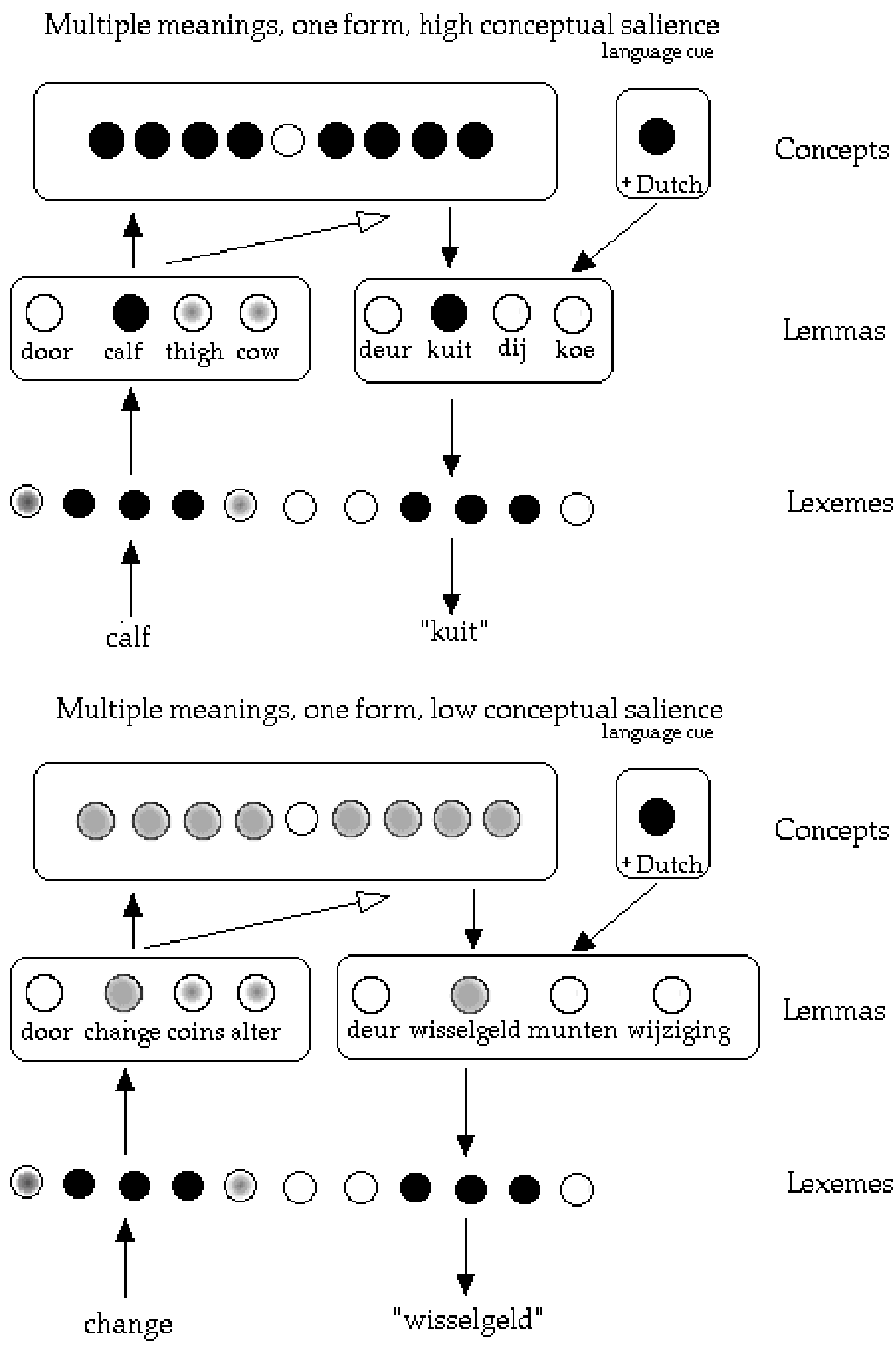


Figure 9. Revised model of conceptual salience and ambiguity effects for words with multiple meanings and one form. The top panel shows words high in conceptual salience and the bottom panel shows words low in conceptual salience.

secondary meanings for low conceptual salience words are essentially nonexistent and therefore performance on these words should be the same as performance for unambiguous low conceptual salience words, which is consistent with the data.

All translations of a word with multiple forms are activated, and this activation makes the form selection process take longer. Note that this effect is fairly constant across conditions, which is consistent with the idea that this simply adds time to the translation process but does not interact with conceptual salience.

Finally, the finding that cognates do not show a difference between the number of meanings and number of forms is most likely due to their more complete overlap across languages at the form level. Perhaps the form overlap in combination with the meaning overlap allows for the form resolution process to occur more quickly than in the noncognate case. Certainly, the stimulus word itself can serve as a memory cue to the translation. Why do meanings and forms have a similar effect on cognate translation? Because when there is a cognate with multiple meanings, only one of these meanings corresponds to the cognate translation. Therefore, meaning resolution may proceed more quickly in these cases because the form in the target language may activate the meaning (e.g., via feedback).

The results of this experiment are important because they show that ambiguity in the form of multiple forms and multiple meanings influence the translation process. This is especially interesting in light of the paucity of research on this topic in the literature. Based on the within-language literature, there is reason to believe that all meanings of a word are activated, even in context, and therefore these effects are relevant to normal language processing. Furthermore, the existence of multiple ways to translate a word most likely will be difficult for learners of a second language, especially when there is a distinction in the L2 that is not present in L1 (e.g., MacWhinney, in press). For example, English has only one word to express the verb ‘to be’, whereas Spanish has two verbs to express that concept, “ser” and “estar”. It is quite difficult for learners of Spanish to learn the appropriate contexts in which to use ser and estar. If the same is true for the lexical level as the syntactic level, the existence of multiple translations for a word will not only affect translation performance in an on-line task, but also the process of language learning. It is also probable that language production will be influenced by these non-native distinctions because learners often must rely on rules rather than native-like intuitions.

Do the results of this experiment reflect a general aspect of conceptual representation or rather aspects of conceptual representation that are specific to bilinguals? Experiment 2 examined the effects of ambiguity in the form of multiple meanings and conceptual salience on a within-language comprehension task. The results of that experiment were then compared to those of the present experiment to determine whether the results found here are specific to bilinguals. Finally, Experiment 3 will examine the time course of the activation of alternative translations using a time course manipulation of the translation task.

Chapter 3

Experiment 2: Within-language Lexical Decision

The results of Experiment 1 showed that translation performance is influenced by both conceptual salience and ambiguity, and that these factors interact with cognate status. Furthermore, the source of the ambiguity, the lexical level or the conceptual level, determines the precise effect on translation performance. Experiment 2 was performed to determine whether the interaction observed in Experiment 1 was due to a general property of meaning representation, or rather to a property of meaning representation that is specific to bilingual language processing. Conceptual salience and ambiguity in terms of number of meanings have been shown to influence lexical decision latencies in the past (e.g., James, 1975; Rubenstein, Garfield, & Millikan, 1970), however the two factors have not been examined simultaneously in the past (with one exception, see below). If the interaction observed in translation is the result of a general characteristic of the language system, then we would expect the same results to be obtained in a within-language task. However, if the results were due to a specific manifestation of conceptual salience and ambiguity for cross-language processing, then we would not expect to find the same pattern of results.

Within-Language Ambiguity Effects

In everyday communication, comprehenders are faced with lexical ambiguities or words that are reliant upon context to indicate their meaning. A classic example is the English word “bug” which has several possible interpretations. It could mean an electronic listening device used by private investigators or an insect, among other things. Clearly, context must be used to determine the intended meaning, however, there is evidence showing that all meanings of an ambiguous word are activated, at least briefly, regardless of context (Onifer & Swinney, 1981; Seidenberg et al., 1982; Swinney, 1979; Tanenhaus et al., 1979). An exception to this is when there is highly constraining sentence context that predicts the most common meaning of the ambiguity, in which case only the most common meaning of an ambiguous word will be activated (Tabossi, 1988). However, everyday language is not likely to be comprised solely of highly constraining context.

Past research on number of meanings (NOM) has, in most cases, suggested that words with more than one meaning have a speed advantage over words with one meaning in lexical decision, a task in which participants indicate whether letter strings are real words (e.g., Hino & Lupker, 1996; Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Millis & Button, 1989; Rubenstein et al., 1970; Rubenstein, Lewis, & Rubenstein, 1971). In contrast, some research suggests that words with multiple meanings are *not* processed more quickly than words with one meaning in lexical decision (e.g., Forster & Bednall, 1976; Gernsbacher, 1984). Thus, although the results in the NOM literature have been inconsistent, words with many meanings are never responded to more slowly than words with few meanings. The stimulus selection procedure employed in past NOM studies will be considered below as a possible explanation for the discrepant results.

The within-language number of meanings variable differs slightly from its cross-language counterpart. The within-language variable is a direct measure of the number of meanings of a word. The cross-language variable is a direct measure of the number of meanings a word has *that have consequences for translation*, but an indirect measure of the number of meanings a word has. Therefore, the acronym NOM will be used to refer to only the within-language variable.

Note that in Experiment 1, ambiguous words were responded to more slowly and with less accuracy than unambiguous words. This disadvantage for ambiguous words in translation contrasts with results from the lexical decision task in which ambiguous words are often responded to more quickly than unambiguous words. This difference is likely to be due to a difference in the requirements of the two tasks. Translation requires one word to be selected for output, and when a single node must be chosen from among many, the resulting active competition slows the translation process.

In contrast, lexical decision is a comprehension task and the activation of a lexical entry or meaning is needed to make a lexical decision (e.g., Forster & Chambers, 1973). Therefore, the super-threshold activation of *any* node will suffice, which results in *passive competition* among the nodes. It is assumed that each meaning of an ambiguous word is represented as a separate node in memory (e.g., Rubenstein et al., 1970, Rubenstein et al., 1971). In a passive-competition model, each possible node accrues activation independently and then the first to cross the threshold determines RT (e.g., Raab, 1962; Rubenstein et al., 1970). Passive competition therefore gives words with multiple nodes more chances to cross the threshold quickly (e.g., Azuma & Van Orden, 1997). Therefore, on average, words with multiple nodes will be responded to more quickly because they have multiple opportunities to be responded to quickly (i.e., due to statistical facilitation), whereas words with one node will have only one chance.

Because of the differences between the translation and lexical decision tasks, it is possible that the effects of conceptual salience and ambiguity on these two tasks will differ.

Indeed, a recent study shows that the typical lexical decision advantage for ambiguous words turns to a disadvantage when the specific meaning must be understood, as in a reading task (Piercey & Joordens, 2000). However, it is difficult to find within and cross-language tasks that have similar requirements, as will be discussed below.

Within-Language Conceptual Salience Effects

As mentioned in the general introduction, conceptual salience (concreteness and context availability) has been shown to influence lexical access in within-language tasks. In the lexical decision task in particular, words high in conceptual salience are generally responded to more quickly than words low in conceptual salience (e.g., James, 1975; Kroll & Merves, 1986; Schwanenflugel et al., 1988; Schwanenflugel & Shoben, 1983).

Although both ambiguity and conceptual salience have been shown to influence lexical access, these two variables have not been studied simultaneously.[14] However, Jastrzembski (1981) examined the number of meanings of the stimuli used in a study that investigated concreteness effects (Ellis & Shepherd, 1974) and showed that the concrete words had slightly more meanings than the abstract words (24.0 versus 18.9, as determined using an unabridged dictionary). Jastrzembski concluded that the concreteness effect reported by Ellis and Shepherd may actually have been a NOM effect. This confounding of

---

14 One exception is a study reported by Rubenstein et al. (1970), in which an interaction between concreteness and NOM was found. However, in that study words were classified as abstract if they had abstract *and* concrete meanings, and no words with only multiple abstract meanings were included, therefore it is not clear how these results relate to performance on ambiguous words with abstract meanings.

concreteness and ambiguity is similar to the one reported by Schönpflug (1997) however, in that case words higher in concreteness were lower in ambiguity.

A possible explanation for the discrepant results in the polysemy literature is that conceptual salience has not been taken into account in the majority of these studies. Furthermore, it is possible that in some past studies, words low and high in conceptual salience have been differentially distributed across the categories of words with few versus many meanings, and that NOM effects were sometimes eliminated by conceptual salience effects (e.g., if high salience words with few meanings were compared to low salience words with many meanings).

To test the reliability of NOM effects and to examine the independent effects of NOM and conceptual salience, these factors were examined simultaneously with a lexical decision task. The data from the translation task used in Experiment 1 showed that the words with multiple meanings were translated more slowly than unambiguous words only for noncognates high in conceptual salience. If these effects are not specific to cross-language processing and the particular requirements of the translation task, we would expect to find an analogous interaction in this experiment. In particular, an effect of multiple meanings (here, an advantage for words with multiple meanings) only for words high in conceptual salience would be expected.

Method

Participants

One hundred and twelve right-handed native speakers of English participated in this experiment. They received extra credit towards a Psychology course for their participation.

Stimuli

The critical stimuli used in this experiment were 400 words that differed in their concreteness and ambiguity, measured in terms of number of meanings (see the section on number of meanings norms below). The printed frequency for each word per million was taken from Francis and Kucera (1982). The words were divided into five versions; participants were randomly assigned to one of the versions. In addition, an equal number of orthographically permissible nonwords that were matched to the target words in terms of length (number of letters) were included for the "no" trials. With several exceptions to accommodate length differences across lists, the same nonwords were used in all versions of the task. See Table 19 for examples of the stimuli.

Table 19

Sample Stimuli from Experiment 2

Real Words		
	Conceptual Salience	
Number of Meanings	Low	High
Few	anguish	apple
Many	youth	leg
Nonwords		
	sprink	yinger
	diger	mattle

Concreteness and Context Availability Ratings. Concreteness and context availability ratings were obtained from six and ten native English speaking students at The Pennsylvania State University, respectively. They received extra credit towards a psychology course for their participation. The instructions used for the concreteness and context availability ratings were adapted from Spreen and Schulz (1966; see Appendix E) and Schwanenflugel et al. (1988; see Appendix F), respectively.

Number of Meanings Norms. The number of meanings norms were obtained from fifty native speakers of English who were students at The Pennsylvania State University. They received extra credit towards a psychology course for their participation. The words were divided into five lists, each list was assigned an approximately equal number of words high and low in conceptual salience. The words were printed in random order in booklets and participants wrote all *distinct* meanings they thought of for each word in one of the five lists. Their responses were coded for accuracy using the information for each word listed in the American Heritage Electronic Dictionary (1992).

Number of meanings norms were gathered using the average meaning metric (e.g., Millis & Button, 1989). The number of noun, verb, adjective, and adverb meanings listed for each word by each participant were calculated. The average number of meanings for each meaning type across participants was then determined. The sum number of meanings across all syntactic categories (i.e., the sum of the averages) was taken to indicate the total number of meanings of a word, and will be referred to as sum meanings. This measure of number of meanings is distinct from the measure used to obtain number of translations norms because the number of meanings norms required that the participants list separately what they considered to be *distinct* meanings of a word. It would be difficult to distinguish meanings if each participant only listed the first meaning that came to mind.

Lexical Decision Procedure

All participants were tested individually and were given verbal instructions by the experimenter, followed by additional instructions on the computer screen. A practice block with 12 trials (6 real words) was given prior to the critical block. Participants were presented with the stimuli one word at a time, at the center of the computer screen. Each stimulus was preceded by a fixation point that remained on the screen for 500 ms. There was an ISI of 100 ms, after which the stimulus appeared and remained on the screen until the participant responded. Participants were instructed to respond to each item by pressing one button to indicate that the stimulus was a real word in English and another button if it was not. Reaction time in ms was recorded by the computer program from the onset of stimulus presentation to the response. There was an ITI of 250 ms. The PsyScope computer program was used to present the stimuli in a random order and to record reaction time and accuracy (Cohen et al., 1993).

Results

Data Trimming

Data from two participants were lost due to equipment failures. Data from 35 participants who did not respond accurately at least 90% of the time for both real and nonwords were removed from the analyses.[15] The final analyses included data from 75

---

[15] The data from these participants were excluded on the basis that they may not have paid sufficient attention to the task. This conclusion was drawn due to their relatively low accuracy on the nonwords (86 %). In addition, the overall printed word frequency of the real word stimuli was rather high, 86 occurrences per million, and even the words low in frequency are fairly familiar (e.g., trauma, malice). For these reasons, the data from these participants were replaced by participants who were tested subsequently and who met the accuracy criteria.

participants (15 for each of the five stimulus lists). Reaction times that were faster than 300 ms or slower than 3000 ms were removed from the analyses, resulting in the removal of 0.10 % of the data. In addition, reaction times that were 2.5 *SD*s above or below the participant's mean RT were excluded from the analyses and were treated as missing values. These criteria were applied to the words and nonwords separately and resulted in the removal of 2.7 and 2.9 % of the data for the words and nonwords, respectively. The data from one word, "organdy", was removed because it was never responded to accurately.

Preliminary Analyses

In this experiment as in Experiment 1, the regression approach to data analysis was taken, and the effects of covariates (i.e., length and frequency) were removed. Furthermore, this approach allowed for the use of a less artificial stimulus set than would be possible using an analysis of variance approach.

Correlations. The correlations between the predictor variables were calculated to determine whether there were significant intercorrelations among the variables (see Table 20). The results show that there were several significant intercorrelations. The concreteness ratings were correlated with the context availability ratings, the concreteness by NOM interaction and the context availability by NOM interaction. Context availability ratings were also correlated with the sum meanings and the concreteness by NOM and context availability by NOM interaction terms. Sum meanings was also correlated with the concreteness by NOM and context availability by NOM interaction terms. The correlations of the predictors with the interaction terms are expected (Aiken & West, 1991) because the interaction terms

Table 20

Intercorrelations Between Factors for the Reaction Time Analysis of Experiment 2

Factor	1	2	3	4	5
1. Concreteness Rating	--	.70**	-.04	.68**	.17**
2. Context Availability Rating		--	.13**	.53**	.39**
3. Sum meanings			--	.40**	.59**
4. Concreteness X Sum Meanings				--	.77**
5. CA X Sum Meanings					--

Note. ** $p < .01$

are made up of the product of two factors; the interaction terms therefore will be correlated with the two factors from which they are derived. The correlation between concreteness and context availability makes it impossible to determine the independent contribution of these predictors. Therefore, as in Experiment 1, concreteness and context availability were used as a set (conceptual salience) to predict performance on the lexical decision task.

Furthermore, the standardized predicted values and residuals from the reaction time analysis were saved and then plotted as scatterplots to verify the validity of using linear regression analyses (Cohen & Cohen, 1983; see Appendix G). These plots suggested that there were no problems of heteroscedasticity or curvilinearity; linear regression analyses were therefore used to analyze these data.

Reaction Time and Accuracy Data

Real word trials. A hierarchical regression analysis was used to analyze the data from correct trials on real words only. The stimulus length and logarithmic transform of word frequency served as covariates. Concreteness and context availability ratings, which have more consistently predicted lexical decision latencies than ambiguity in the past, were entered on the second step. Then, the sum number of meanings were entered. Finally, the interaction between the conceptual salience ratings and the number of meanings was entered. The dependent variable was the mean reaction time. The change in the proportion of variance (ΔR^2) accounted for by each step in the analysis as well as the associated semipartial correlation coefficients (*sr*) were evaluated (see Table 21).

The results of the hierarchical regression analysis showed that a significantly greater proportion of variance was accounted for when conceptual salience and the sum number of meanings were in the model (step 3) than when conceptual salience alone was in the model (step 2). The interaction terms did not account for additional variance. The length and frequency of the stimulus were significant predictors of reaction time, as expected. The effect of conceptual salience was significant, such that words high in conceptual salience were responded to more quickly than words low in conceptual salience. Furthermore, the effect of number of meanings was also significant, such that the more meanings a word had, the more quickly it was identified as a word. High and low levels of conceptual salience ($M \pm 1$ *SD*) were entered into the regression equation with low and high levels of sum meanings (1 and 2, respectively). The resulting predicted lexical decision latencies are given in Table 22.

Table 21

Overview of the Hierarchical Regression Analysis for Variables Predicting Reaction Time in the Lexical Decision Task

Step	Variable	β	*sr*
1	Stimulus length	.34*	.33*
	Log frequency	-.19*	-.19*
2	Stimulus length	.23*	.22*
	Log frequency	-.15*	-.15*
	Concreteness rating	.29*	.20*
	Context availability rating	-.59*	-.40*
3	Stimulus length	.21*	.19*
	Log frequency	-.15*	-.15*
	Concreteness rating	.25*	.17*
	Context availability rating	-.56*	-.37*
	Sum number of meanings	-.16*	-.15*
4	Stimulus length	.21*	.19*
	Log frequency	-.15*	-.15*
	Concreteness rating	.29	.08
	Context availability rating	-.58*	-.28*
	Sum number of meanings	-.18*	-.14*
	Concreteness x sum meanings	-.04	-.01
	Context availability x sum meanings	.07	.02

Note. Step 1 $R^2\Delta = .177$, F (2, 404) = 43.37, $p < .01$; Step 2 $R^2\Delta = .172$, F (2, 402) = 53.09, $p < .01$; Step 3 $R^2\Delta = .023$, F (1, 401) = 14.89, $p < .01$; Step 4 $R^2\Delta$= .001, F (2, 399) = .28, $p > .10$.

* $p < .01$.

Table 22

Estimated Reaction Times as a Function of Number of Meanings and Conceptual Salience

	Number of Meanings	
Conceptual Salience	One	Two
Low	622	579
High	580	536

Because of a distributional violation regarding the assumption of homoscedasticity and a limited range of variance, a regression analysis was not used to evaluate the predictors of accuracy. On average, the words were responded to with 94.9 % accuracy (*SD* = 9.4), and the range was from zero to one hundred per cent. Only one word was never responded to accurately. With this word excluded, the mean accuracy was 95.1 % (*SD* = 8.2) and the range was from 33 to 100 %.

Effects of Word Status. The average nonword and real word data across participants and items were compared using t-tests to evaluate word status effects on reaction time and accuracy. As in past research, nonwords were responded to more slowly than real words for both the analysis by participants [678 versus 589 ms; t (74) = 11, $p < .01$] and by items [679 versus 592 ms; t (1031) = -20.07, $p < .01$] (e.g., Kroll & Merves, 1986; Rubenstein et al., 1970; Schwanenflugel et al., 1988). Nonwords were also responded to less accurately than real words according to the analysis by participants [93.7 versus 94.6 %; t (74) = 3, $p < .01$] and by items [94.6 versus 93.7 %; t (1032) = 1.65, $p < .05$].

Discussion

Although an interaction between conceptual salience and ambiguity has been observed across languages in translation (the present Experiment 1), no such interaction was observed between conceptual salience and number of meanings in this lexical decision task. The results of this experiment show that when these two factors are examined simultaneously, they both significantly predict lexical decision performance.

The main effects of ambiguity and conceptual salience are consistent with the models of these effects that have been described in the literature. The finding that ambiguous words are responded to more quickly than unambiguous words is consistent with a passive competition model like the ones proposed by Rubenstein et al. (1970) and Azuma and Van Orden (1997). These models assume that words with multiple meanings have a statistical advantage over words with one meaning in comprehension tasks.

The finding that words high in conceptual salience are responded to more quickly than words low in conceptual salience is consistent with the general models of these effects that were described in the general introduction (e.g., Paivio, 1971; Schwanenflugel et al., 1988; Schwanenflugel & Shoben, 1983). Paivio's explanation of these effects is based on the existence of an imaginal code associated with words high in conceptual salience. The context-availability hypothesis (e.g., Schwanenflugel et al., 1988) explains these effects in terms of the greater availability of internally-generated context for words high in conceptual salience than words low in conceptual salience.

There are several additional implications of these findings. First, Jastrzembski's (1981) hypothesis that concreteness effects are actually NOM effects is not supported by these data. Although these two factors may have been confounded in past research, concreteness and NOM appear to have direct and independent effects on the process of lexical access. Furthermore, the findings of this study suggest that if ambiguous words low in conceptual salience were compared to unambiguous words high in conceptual salience in a NOM study, an apparent null effect could be observed. This finding is important because past research has not examined both conceptual salience and ambiguity simultaneously and the effects of one variable may be masked by the other when confounds are present in the materials.

Finally, the interaction observed between number of translations and conceptual salience in bilingual translation does not extend to the within-language domain. There are several possible explanations for this finding. First, it is possible that distinct processes govern within and cross-language processing. However, this is unlikely to be true especially for adult bilinguals whose language processing systems and conceptual representations are well formed prior to the acquisition of the second language. Unfortunately, task differences make this conclusion premature.

A more plausible explanation is that the interaction of various factors that affect language processing will vary with the nature of the task. For example, the task used in the bilingual domain was a production task, whereas the task used in this experiment was a comprehension task. Production tasks force the speaker to select a single node in memory for production and therefore results in competition among nodes in memory. In contrast, comprehension tasks allow for a more shallow level of analysis at which any amount of activation of a node above some threshold will be sufficient. Thus, the difference between the findings in the within and cross-language experiments are most likely due to differences in task demands.

Differences due to task requirements suggest that the goal of the task must be taken into account when predicting performance. In addition, tasks will differ in the level of processing that they require. The findings of a recent study (Piercey & Joordens, 2000) suggest that the processing of ambiguous words goes from a fast, efficient process during which a blend of all meanings is activated. Then, if needed, this blend state will move closer to one of the interpretations of the word. This latter stage is hypothesized to be slower and less efficient than the first. Furthermore, the need to understand a single meaning of the word (as in reading), will lead to the engagement of this latter process.

The differences between lexical decision and translation make it somewhat difficult to determine whether the lack of an interaction between conceptual salience and ambiguity is due to differences in the task or the cross-language versus within-language nature of the tasks. Unfortunately, there is not a task that is comparable to translation in the within-language domain, which is why lexical decision was used in the present experiment. The within-language task used most often in production experiments is picture naming, however that task requires that the stimuli be pictureable nouns and therefore conceptual salience cannot be manipulated.

In future research, it will be possible to examine the influence of conceptual salience and ambiguity on cross-language performance using a comprehension task, however. Lexical decision studies have been done both with one of the two languages, and also using a general method in which a word in any language should be responded to positively. However, it is not clear that the within and cross-language lexical decision tasks are identical, thus caution should be taken in comparing results across these tasks.

Another possible avenue to explore these effects would be to examine animacy effects (differences on living and nonliving kinds; e.g., Sholl, 1995). A task such as picture naming could be used with animacy varied. The difference between animate and inanimate kinds on translation performance is roughly analogous to differences in conceptual salience, although the sources of the effects may be distinct.

In conclusion, the results of this experiment suggest that words higher in conceptual salience are responded to more quickly than words lower in rated conceptual salience on a within-language lexical decision task. Furthermore, words are responded to more quickly when they have more meanings. Finally, there was no interaction between these factors in the present study, which may be the result of the differences in task requirements between translation and lexical decision.

Chapter 4

Experiment 3: Time Course of Translation Production

The results of Experiment 1 showed that translation was slower and more error-prone the more ambiguous the words were. Furthermore, ambiguity interacted with conceptual salience and cognate status. The effects of multiple meanings were reduced for cognates and enhanced for concrete words, whereas the effects of multiple translations were relatively constant across word types. The results of that experiment suggest that the existence of multiple meanings and forms do indeed influence translation performance. However, the source of this effect is still somewhat unclear. Do all meanings of a word become fully activated? Does more than one translation (form) become phonologically encoded (i.e., readied for production)? The model presented in the conclusion of Experiment 1 makes certain predictions regarding the time course of translation; these predictions were examined in the present experiment.

Furthermore, the results of Experiment 1 suggest that there are effects of both the number of meanings and number of forms the bilingual must choose among. However, these two factors may affect performance at different points in time because they operate at distinct levels of the system. Past research on language production suggests that production can be divided into two main components: *lemma selection* and *word form retrieval*. A *lemma* is a representation that includes both semantic and syntactic information for a word, but not its phonological word form (e.g., Levelt et al., 1991). During lemma selection, a semantically-appropriate item is selected and is specified with respect to syntax but not phonology. During word form (lexeme) retrieval, phonological specification occurs (i.e., when the phonological features of the to-be-produced word become activated).

Although models of language production tend to agree that language production involves lemma selection and word form retrieval, they differ in regard to the interaction of these stages. Some models propose that lemma selection must complete before word form retrieval occurs, and therefore only one word will become phonologically encoded, namely the word that is eventually produced (e.g., discrete two-stage models, Levelt et al., 1991). Other models propose that these two stages overlap in time and therefore more than one word can become phonologically encoded (e.g., interactive activation models, Dell & O'Seaghdha, 1991; cascade models, Peterson & Savoy, 1998).

The model presented by Tokowicz and Kroll was a form of interactive-activation model in that the processing at one level did not have to terminate before this information could spread to the next level. The revised model presented in the conclusion of Experiment 1 is a more serial model in which a meaning is first selected and then the forms that correspond to that meaning are activated. This change was made because the predictions made by the Tokowicz and Kroll model were not supported by the results of Experiment 1.

Time Course Task

In the present experiment, a translation variant of the picture naming task used by Peterson and Savoy (1998) was employed. In that study, the goal was to understand whether multiple names for line drawings that had near-synonymous names (e.g., “couch” and “sofa”) became phonologically encoded during picture naming. There were two trial types: word naming (critical) trials and picture naming (filler) trials. On the picture naming trials, first a picture would appear on the screen, followed by a cue (?) that indicated that the participant should name the picture. On the critical trials, a word naming task was performed instead. On these trials, following the picture, a word appeared instead of the picture naming cue, and the participant was instructed to name the word aloud. These words were related or unrelated to one of the picture's two names (see Table 23 for sample stimuli).

Table 23

Sample Stimuli from the Picture Naming Task used by Peterson and Savoy (1998)

Picture Name		Phonological Target	
Dominant	Secondary	Dominant	Secondary
Equivocal Pictures			
Couch	Sofa	Count	Soda
Plate	Dish	Place	Ditch
Unequivocal Pictures			
Drum	--	Drug	--
Grapes		Grades	

Note. The equivocal pictures are those with multiple names, and the unequivocal pictures are those with only one name, as determined in a norming task.

Peterson and Savoy measured the difference in reaction time due to the relation between the word and the picture names. If the words related to the secondary or less dominant name for the picture were responded to either more or less quickly than words unrelated to these names, it could be concluded that the alternate name had become phonologically encoded. Peterson and Savoy manipulated the SOA and found no effect of relatedness at the earliest SOAs, priming (facilitation) for words related to both the dominant and secondary names at the intermediate SOAs, and priming only for the dominant name at the longest SOA. They concluded that there is phonological encoding of secondary names of pictures that begins early in the picture naming process and remains until the dominant name for the picture is chosen. At that point, the activation of the secondary name declines.

In two further experiments, Peterson and Savoy obtained evidence that although semantically-related concepts also become activated during the picture naming process (e.g., “bed” for the couch/sofa picture), these concepts are not phonologically encoded. This finding suggests that only close alternatives for production become encoded during language processing.

The present experiment was an extension of the Peterson and Savoy study to the translation task. Although picture naming is the task used most often to develop and evaluate models of bilingual language production, because picture naming is limited to the domain of pictureable objects, an adaptation of this task was used in the present study. Furthermore, because multiple synonyms (forms) are somewhat analogous to having multiple names for pictures in that there are multiple options for output when provided with the same input, it was expected that this task would show whether or not multiple translations become phonologically encoded during the translation task.

Translation Variant of the Time Course Task

In this experiment, participants were presented with words to translate. On the "naming" (critical) trials, a word was presented, followed by a different word below the stimulus word (the "probe"), that indicated that the participant was to read this word aloud rather than translate the first stimulus. On the "translation" (filler) trials, a question mark was presented just below the stimulus word, that indicated to the participant that he or she was to translate that word. More translation than naming trials were included to encourage the participants to begin translating the words on all trials. The naming trials are the critical trials because it is on these trials that the effect of relatedness (between the translation of the stimulus and the probe word) is measured.

There were two versions of the task based on language of production: English or Dutch. These two conditions roughly correspond to L1 to L2 and L2 to L1 translation, respectively, because the majority of the trials were translation trials. Because the critical stimuli used in the two language-of-production conditions were different, it was not possible to directly compare the two directions of translation. Instead, the patterns of results from the two language-of-production conditions were compared.

The critical words used on the naming trials were phonologically related (or unrelated) to one of the translations of the word. Furthermore, two SOA conditions were used to track the time course of the activation of alternate translations. See Figure 10 for a graphic illustration of the task, and Table 24 for sample stimuli.

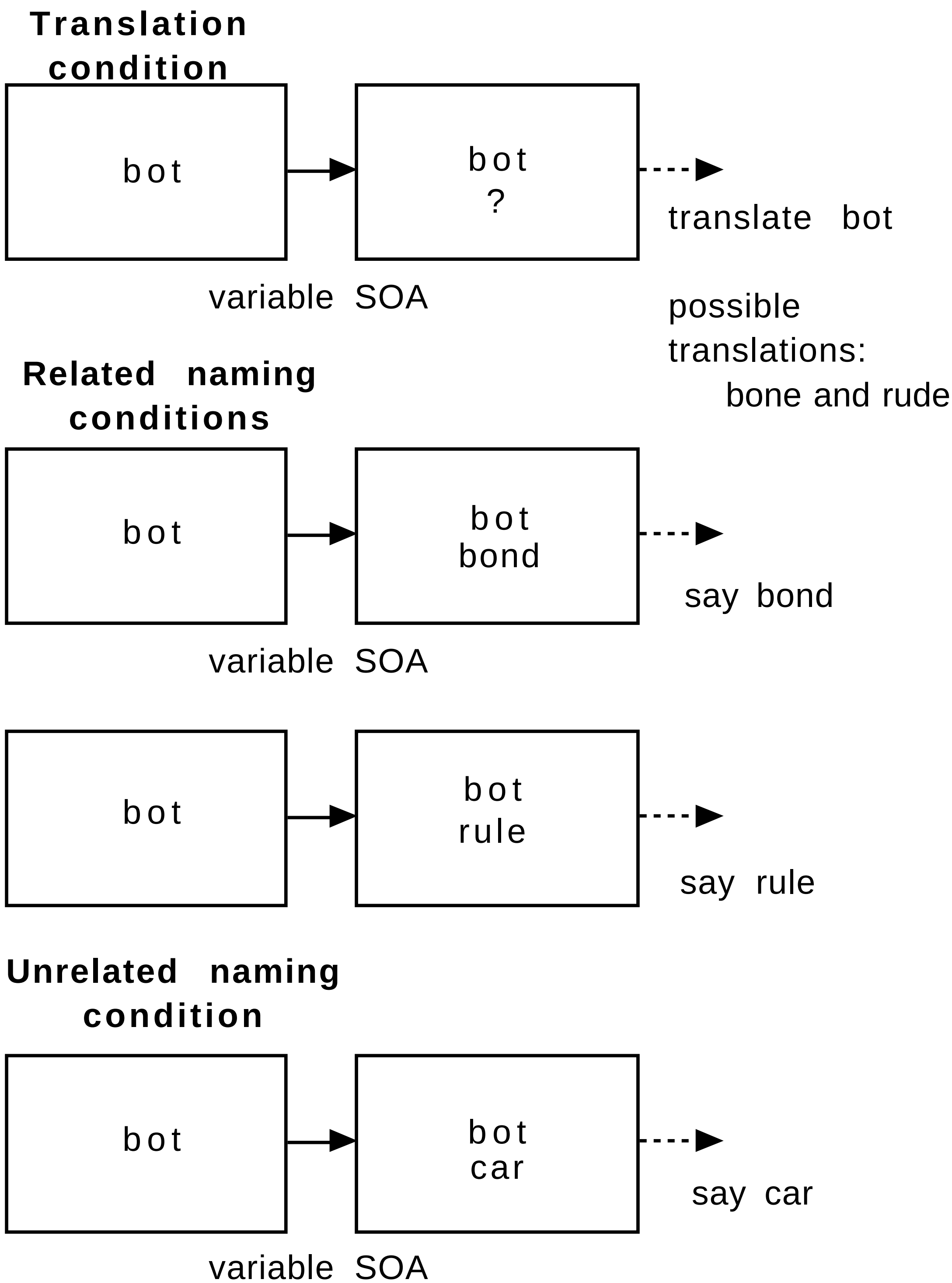


Figure 10. Graphic illustration of the task used in Experiment 3 (phonological condition).

Table 24
Sample Stimuli from Experiment 3

Condition	Stimulus	Translation	Related target	Unrelated target
Phonological Condition: English Production				
Form 1	herfst	autumn	autonomy	cocktail
Form 2	herfst	fall	tall	shift
Meaning 1	boodschap	message	messy	donor
Meaning 2	boodschap	errand	error	pitch
Phonological Condition: Dutch Production				
Form 1	road	weg	zeg	hok
			participation	cage
Form 2	road	straat	straal	hoogst
			beam	highest
Meaning 1	favour	gunst	kunst	vorig
		(privilege)	art	former
Meaning 2	favour	voorkeur	voordeur	liberaal
		(preference)	front door	liberal
Semantic Condition: English Production				
Form	ruzie	quarrel	argument	chapter
Meaning 1	bal	ball	party	early
Meaning 2	bal	ball	globe	ivory

Note. The English translation is given below the Dutch targets.

The two SOA conditions were 300 and 600 ms. These two values were chosen to correspond roughly to approximately one third and two thirds of the time it typically takes to translate a word. Further, these two values were adapted from those used in a previous study with a similar method in which differences between the two conditions were found, as had been expected (Miller, 1997).

Furthermore, an additional condition was added in which the critical words on the naming trials were semantically related or unrelated to one of the translations of the word (see Table 24 for examples). For words with multiple form translations, semantically-related words are related to both translations because the translations are near synonyms. In contrast, for words with multiple meaning translations, one related and unrelated word was chosen for each translation of the stimulus. This condition will be referred to as the semantic condition; this condition was only performed from Dutch to English and the same SOA values were used in this condition as in the phonological conditions.

Hypotheses

Any significant difference in response time between the related and unrelated conditions suggests that the particular translation of a word has been phonologically encoded. To eliminate other factors as alternative explanations for this result, each related probe was matched to its corresponding unrelated probe as closely as possible in terms of length and frequency.

The model presented in the conclusion of Experiment 1 can be used to derive predictions for this task. According to that model, one must first select a meaning of a word and then the translations for that particular meaning will become activated. Figure 11 shows the predicted activation patterns for multiple meanings and multiple forms under this assumption.

Any difference between the related and unrelated conditions is sufficient to indicate activation of an alternate meaning or translation; however, either inhibition (related faster than unrelated) or facilitation (related slower than unrelated) may be obtained.[16] Because the words with one translation can only be translated in one manner, we should observe large relatedness effects for these words. This is because words with one translation will be started to get translated in the same way therefore there should be a large effect of the relatedness of the naming probe to the translation. There would be a smaller relatedness effect if the participant did not begin to translate the word in the anticipated manner (i.e., when there are multiple translations possible).

In the multiple forms case, there should be activation of all forms but when one form has been selected, the activation of the other form(s) will decrease. Therefore, there should initially be activation of all forms, and this would lead to phonological encoding of multiple forms.

The hypotheses for words with multiple meaning translations are based on the assumption that one must first choose a meaning and then select among the forms for that

---

[16] Note that no model of language production predicts facilitation or inhibition in particular. This is because any difference is considered sufficient to show that there is an effect of relatedness.

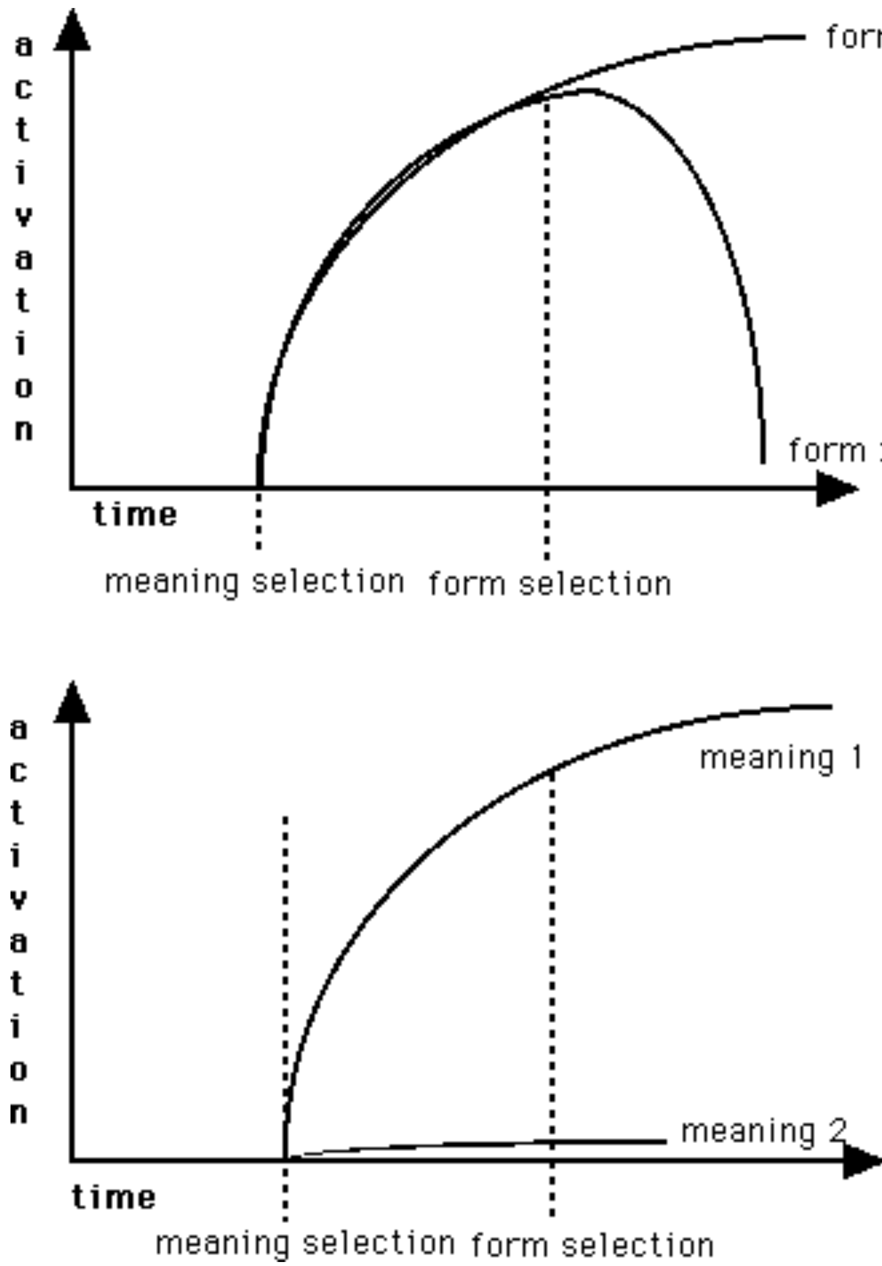


Figure 11. Predicted activation functions from a serial two-stage processing model. The top graph represents a word with multiple form translations and the bottom graph represents a word with multiple meaning translations.

meaning. Given this assumption, the initial access of all meanings will lead eventually to only one meaning being highly activated, and it is this meaning that will get phonologically encoded. Thus, there should only be phonological encoding of the translation for the active meaning.

The difference between the two types of words (with multiple forms and multiple meanings) is that the multiple forms should show activation of both forms whereas the multiple meanings should only show activation of one. In this task we would expect the activation of one versus two translations to appear as smaller versus larger priming effects in the two cases.

During meaning selection, activation of both meanings of a multiple meaning word should be observed. After meaning selection has taken place, activation of only the selected meaning should be observed.

Method

Participants

The participants in the this experiment were 48 Dutch-English bilinguals who were students at the University of Nijmegen (16 per condition). They were paid or given extra credit towards their studies in psychology for their participation. Their language history questionnaire data are shown in Table 25 by condition.

Table 25

Language History Questionnaire Data from the Participants of Experiment 3 by Condition

	Condition		
Measure	Phonological: English Production	Phonological: Dutch Production	Semantic: English Production
Age (years)	22.5 (3.8)	22.3 (2.9)	21.9 (2.0)
Age Began L2 (years)	8.7 (3.7)	10.3 (2.0)	9.3 (2.6)
Time Studied L2 (years)	11.0 (3.8)	9.9 (3.0)	11.4 (2.6)
L2 Immersion Experience (months)	1.9 (4.7)	2.5 (4.5)	3.6 (6.6)
L2 Reading Ability	8.4 (1.1)	8.1 (1.2)	8.7 (.7)
L2 Writing Ability	7.5 (1.3)	7.1 (1.2)	7.6 (.6)
L2 Conversation Ability	7.7 (1.2)	7.7 (1.3)	7.6 (1.1)
L2 Speech Comprehension Ability	8.7 (.9)	8.4 (1.2)	9.0 (.5)

Note. Standard deviations are given in parentheses. Reading, writing, conversational, and speech comprehension ability were rated on a 10-point scale where 1 indicated the lowest level of ability and 10 indicated the highest level of ability. The participants in the two phonological conditions did not differ significantly on any of these measures (see Appendix B for additional information on group differences).

Design

In this study, there were different numbers of stimuli in the various conditions due to the different numbers of words in these conditions in the stimulus set and presumably in natural language (see Table 26 for the phonological condition, English production, Table 27 for the phonological condition, Dutch production, and Table 28 for the semantic condition, English production). There were two trial types: naming (critical) and translation (filler). Of the naming trials, there were some stimuli with multiple form translations, some with multiple meaning translations, and some with only one translation. Within each of these conditions, there were two relatedness conditions (phonologically related and unrelated) and two SOA conditions (300 and 600 ms). Furthermore, for words with multiple translations, there were related and unrelated probes for each of the translations of the word (translation 1 and translation 2).

One complication of the results of this experiment is due to the way in which translation number was assigned. Although not completely arbitrary, this factor was not predicted to have an effect because most words had translations of equal probability (i.e., the words had multiple translations that were approximately equally likely to be given as translations of that word during the norming task). This selection procedure was used because these are the conditions under which the secondary or (slightly) less dominant translations would be most likely to be considered candidates for production (e.g., Jescheniak & Schriefers, 1998). The words were assigned translation number based on the words that were part of the original stimulus pair from De Groot (1992). That is, the original translation

Table 26

The Number of Stimuli in the Phonological Condition of Experiment 3: Production in English

Production in English											
Naming Trials											
Multiple Form Translations				Multiple Meaning Translations				One Translation			
Related/Unrelated				Related/Unrelated				Related/Unrelated			
Translation 1		Translation 2		Translation 1		Translation 2					
SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600
6	6	6	6	3	3	3	3	12	12	12	12
Translation Trials											
Multiple Translations						One Translation					
SOA 300			SOA 600			SOA 300			SOA 600		
23			22			127			128		

Note. The same number of stimuli were in the related and unrelated conditions. Thus, there were 6 words related to a word with multiple form translation's first translation at the 300 ms SOA condition, and 6 additional words in the corresponding unrelated condition.

Table 27

The Number of Stimuli in the Phonological Condition of Experiment 3: Production in Dutch

Production in Dutch											
Naming Trials											
Multiple Form Translations				Multiple Meaning Translations				One Translation			
Related/Unrelated				Related/Unrelated				Related/Unrelated			
Translation 1		Translation 2		Translation 1		Translation 2					
SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600
6	6	6	6	3	3	3	3	18	18	18	18

Translation Trials			
Multiple Translations		One Translation	
SOA 300	SOA 600	SOA 300	SOA 600
51	50	173	174

Note. The same number of stimuli were in the related and unrelated conditions. Thus, there were 6 words related to a word with multiple form translation's first translation at the 300 ms SOA condition, and 6 additional words in the corresponding unrelated condition.

Table 28

The Number of Stimuli in the Semantic Condition of Experiment 3: Production in English

Production in English									
Naming Trials									
Multiple Form Translations		Multiple Meaning Translations				One Translation			
Related/Unrelated Both Translations		Related/Unrelated Translation 1		Translation 2		Related/Unrelated			
SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600	SOA 300	SOA 600
12	12	3	3	3	3	18	18	18	18

Translation Trials			
Multiple Translations		One Translation	
SOA 300	SOA 600	SOA 300	SOA 600
28	29	108	108

Note. The same number of stimuli were in the related and unrelated conditions. Thus, there were 12 words related to a word with multiple form translation's first translation at the 300 ms SOA condition, and 12 other words in the corresponding unrelated condition.

given by De Groot was always considered "translation 1" and the alternate translation of the highest dominance (in the case of more than one alternate) was considered "translation 2". It was assumed that this method of assignment would not bias the results because most words had equi-probable translations. However, the dominant translation was labeled as translation 2 on approximately 20 % of the trials in each language-of-production condition. See the Stimulus section for more information regarding the critical items.

There were two language-of-production conditions: English production which consisted of Dutch to English translation and naming words in English, and Dutch production which consisted of English to Dutch translation and naming words in Dutch. Several versions of the stimuli were created so that a given stimulus appeared in one condition with a related word and in another condition with an unrelated word; a given word pair appeared at both SOA conditions in different versions of the experiment. That is, across participants each particular stimulus word was part of every condition (related, unrelated, etc.), however each stimulus was only presented once to a given participant. Participants were randomly assigned to one of the versions of the experiment, and the stimuli were presented in random order.

Stimuli

The stimuli were chosen from the same set of words used in Experiment 1. Translation trials were included as fillers to ensure that the participants prepared to translate each item. Different words served as critical items in the conditions of the experiment because different words had multiple translations in each direction of translation. Thus, the critical stimuli for the English production condition were those words that had multiple translations from Dutch to English. The probe items used on the naming trials were either phonologically related to one of the translations of the stimulus word, or unrelated to any of the translations. In the semantic condition, the probes were semantically related or unrelated. The items in each condition (form/meaning, related/unrelated, 300/600 ms SOA) were matched as closely as possible on word length, frequency, concreteness, and cognate rating. See Table 29 for the properties of the stimuli in the Dutch production condition and Table 30 for the properties of the stimuli in the English production condition. See Table 31 for the properties of the stimuli in the semantic condition.

The logic of the stimulus selection procedure was that the related and unrelated naming targets should be as similar as possible in terms of length and frequency because these factors have been shown to affect naming latencies. Therefore, each related target was matched to its unrelated target in length and frequency (exactly, in most cases). It was not possible to exactly match the words in the form and meaning conditions on these dimensions because there were a limited number of stimuli in these categories. However, it is still possible to compare the overall pattern of the results for these two conditions because the relatedness effect should not be influenced by the factors that were matched between the related and unrelated items (i.e., length and frequency).

Procedure

All participants were tested individually and given verbal instructions by the experimenter, followed by additional instructions on the computer screen. Participants translated aloud visually-presented words either from L1 to L2 and L2 to L1 and named

Table 29

Properties of the Critical Stimuli from Experiment 3: Phonological Condition: English Production

	One Translation		
Measure	Translation	Related	Unrelated
English word length	5.5 (1.7)	5.4 (1.9)	5.4 (1.6)
English word Frequency	134.4 (303.4)	113.5 (206.5)	112.6 (203.6)
	Form Translations		
English word length	6.3 (1.8)	5.9 (1.5)	5.9 (1.7)
English word Frequency	83.3 (109.2)	98.4 (296.8)	97.50 (289.1)
	Meaning Translations		
English word length	6.2 (2.1)	5.7 (1.6)	5.8 (1.7)
English word Frequency	126.9 (183.1)	96.5 (180.6)	96.6 (181.4)
	Alternate Form Translations		
English word length	5.6 (2.3)	5.7 (2.1)	5.7 (1.9)
English word Frequency	144.1 (237.6)	36.9 (55.0)	36.8 (54.9)
	Alternate Meaning Translations		
English word length	6.3 (2.0)	6.0 (1.6)	5.9 (1.6)
English word Frequency	98.3 (113.0)	59.3 (100.9)	67.4 (103.9)

Note. Standard deviations are given in parentheses. There were no significant differences between the words with one, multiple form, and multiple meaning translations on any of these dimensions, *p*s > .10.

Table 30

Properties of the Critical Stimuli from Experiment 3: Phonological Condition: Dutch Production

Measure	Translation	Related	Unrelated
	One Translation		
Dutch word length	5.5 (1.8)	5.5 (1.8)	5.5 (1.7)
Dutch word Frequency	170.8 (549.6)	14.2 (20.6)	14.4 (20.4)
	Form Translations		
Dutch word length	5.5 (1.8)	5.8 (2.0)	5.8 (1.9)
Dutch word Frequency	128.9 (206.3)	35.3 (90.1)	35.7 (90.9)
	Meaning Translations		
Dutch word length	5.0 (1.5)	5.2 (1.88)	5.2 (2.0)
Dutch word Frequency	90.4 (102.5)	32.6 (43.8)	32.6 (43.8)
	Alternate Form Translations		
Dutch word length	6.1 (2.3)	6.0 (2.1)	6.1 (2.1)
Dutch word Frequency	158.8 (297.0)	15.8 (27.5)	15.9 (27.6)
	Alternate Meaning Translations		
Dutch word length	6.3 (2.5)	6.3 (2.5)	6.3 (2.6))
Dutch word Frequency	80.2 (146.2)	22.6 (34.5)	22.7 (34.4)

Note. Standard deviations are given in parentheses. There were no significant differences between the words with one, multiple form, and multiple meaning translations on any of these dimensions, *p*s > .10.

Table 31

Properties of the Critical Stimuli from Experiment 3: Semantic Condition, English Production

	One Translation		
Measure	Translation	Related	Unrelated
English word length	5.7 (1.6)	5.8 (1.7)	5.8 (1.5)
English word Frequency	107.3 (133.2)	123.7 (169.0)	123.1 (165.0)
	Form Translations		
English word length	5.2 (1.5)	6.1 (1.9)	6.1 (1.8)
English word Frequency	64.7 (77.8)	97.2 (115.3)	97.0 (114.7)
	Meaning Translations		
English word length	4.9 (1.7)	5.9 (1.5)	5.8 (1.6)
English word Frequency	116.3 (175.6)	121.0 (127.5)	122.1 (129.9)
	Alternate Meaning Translations		
English word length	5.5 (1.4)	6.3 (2.5)	6.2 (2.3)
English word Frequency	119.7 (219.8)	68.2 (115.2)	68.5 (117.0)

Note. Standard deviations are given in parentheses. There were no significant differences between the words with one, multiple form, and multiple meaning translations on any of these dimensions, *p*s > .10.

words either in L1 or L2. A given participant was only expected to produce words in one language in this experiment. Thus, participants who translated words from L1 to L2 named words in L2 and those who translated words from L2 to L1 named words in L1. No stimulus was ever repeated for a given participant. The order of the trials was randomized by the computer program.

Practice trials were given prior to the critical trials in each direction of translation. Prior to the presentation of each stimulus, a fixation point was presented at the center of the computer screen until the participant initiated the beginning of the trial by pressing a key on the button box. Following the offset of the fixation point, the prime word was presented slightly above the center of the screen. The word remained on the screen until the end of the trial. After either 300 or 600 ms (depending on the SOA condition for that trial), the second stimulus appeared slightly below the center of the screen. This word also remained on the screen until the participant made a verbal response. On the naming trials, the second stimulus was another word to be named aloud. On the translation trials, the second stimulus was a question mark that indicated that the participant was to translate the word. Participants were instructed to respond as quickly and accurately as possible once the second word or question mark appeared on the screen, and to indicate when unaware of the correct translation of the presented word by saying “no” or its Dutch equivalent. Their verbal responses were tape recorded and later coded for accuracy. Reaction time in ms was recorded by the computer voice key from the onset of the second stimulus to the onset of the verbal response. The PsyScope computer program was used to present the stimuli in a random order and record reaction time (Cohen et al., 1993).

Results

Data Trimming

The data from two participants were lost due to equipment failures. The final analyses included data from 46 participants (15 per phonological condition and 16 for the semantic condition). Reaction times that were longer than 3000 ms were removed from the analyses, which resulted in the removal of .2 and 0 % of the data in the Dutch and English production conditions, respectively. Outliers were calculated for the critical (naming) trials only; responses that were 2.5 SDs above or below the mean response time for all naming trials were considered outliers and removed from the analyses. This resulted in the removal of 2.4 and 2.3 % of the trials in the English and Dutch production conditions, respectively. Finally, trials on which the voice key failed were removed from the analyses, which resulted in the removal of 2.9 and 2.8 % of the trials for the English and Dutch production conditions, respectively.

The critical analyses involve the comparison of the naming latencies to determine the effects of number and type of multiple translation. These two types of analyses were performed separately.

Reaction Time Analyses

Effects of Number of Translations on Naming. Are all translations of a word phonologically encoded? The absolute response times for words in the one and more than one translation conditions are given in Table 32 along with difference scores (related minus unrelated RT). Any difference between related and unrelated response times is an indication that phonological encoding occurred.

For the English production condition, there were two significant effects. First, there was an overall effect of number of translations such that words in the one translation condition were named more quickly than words in the multiple translations condition [711 versus 731 ms; F (1, 14) = 7.0, $p < .05$]. In addition, there was a main effect of SOA such that words were named more quickly at the long than the short SOA [701 versus 741 ms; F (1, 14) = 24.7, $p < .01$]. Newman-Keuls tests showed that the only condition that showed a difference between the related and unrelated words was the more than one translation 300 ms SOA condition. This finding suggests that there may have been phonological encoding for this condition. However, there was no effect of relatedness for words with one translation, which suggests that there may be a methodological problem with this experiment. This issue will be discussed further in the discussion section.

Table 32

Reaction Time Data from Experiment 3 by Number of Translations and Condition

	Number of Translations					
	One Translation			More than One Translation		
SOA	Related	Unrelated	Difference	Related	Unrelated	Difference
	Phonological Condition: English Production					
300 ms	727	734	-7	765	737	28
600 ms	687	697	-10	710	712	-2
	Phonological Condition: Dutch Production					
300 ms	736	744	-8	751	746	5
600 ms	696	695	1	700	709	-9
	Semantic Condition: English Production					
300 ms	659	685	-26	672	698	-26
600 ms	612	635	-23	644	634	10

Note. Negative difference scores indicate facilitation and positive scores indicate inhibition.

For the Dutch production condition, the only significant effect was a main effect of SOA which showed that words were named more quickly at the long SOA than at the short SOA [700 versus 744 ms; F (1, 14) = 15.5, p < .01]. No conditions showed a significant relatedness effect.

For the semantic condition, overall, there was a significant effect of number of translations, such that words in the one translation condition were named more quickly than words in the multiple translation conditions [648 versus 662 ms, F (1, 15) = 8.8, $p < .01$]. There was also a significant effect of relatedness, such that words that were related to the stimuli were named more quickly than words that were unrelated [647 versus 663 ms, F (1, 15) = 8.9, $p < .01$]. Finally, there was a significant effect of SOA such that words were named more quickly at the long than the short SOA [631 versus 679 ms, F (1, 15) = 41.9, $p < .01$].

The finding of an overall relatedness effect in the semantic condition is consistent with the idea that alternate conceptual candidates are activated during the translation process (e.g., “bed” for couch/sofa, Peterson & Savoy, 1998). Furthermore, the finding of semantic effects in the absence of the phonological effects in this experiment suggests that the SOA values that were chosen may not have been long enough to demonstrate the phonological effects. This is because semantic effects should occur earlier than phonological effects.

Effects of Type of Multiple Translations on Naming. The absolute reaction times and difference scores for words with multiple forms and multiple meanings are given in Table 33 by condition. Several effects were significant for the phonological English production condition. First, there was a significant effect of translation dominance, such that words in the translation 1 condition were named more quickly than words in the translation 2 condition [725 versus 742 ms; F (1, 14) = 4.6, $p < .05$]. There was also a significant effect of SOA condition, such that words were named more quickly at the long SOA than the short

Table 33

Reaction Time Data from Experiment 3 by Type of Multiple Translations and Condition

	Form						Meaning					
	Translation 1			Translation 2			Translation 1			Translation 2		
SOA	Related	Un-related	Differ-ence	Related	Un-related	Differ-ence	Related	Un-related	Differ-ence	Related	Un-related	Differ-ence
	Phonological Condition: English Production											
300	735	721	14	786	752	34	763	762	1	784	717	67
600	708	705	3	705	692	13	711	698	13	713	785	-72
	Phonological Condition: Dutch Production											
300	756	761	-5	783	739	44	725	708	17	711	770	-59
600	691	729	-38	704	703	1	721	690	31	688	698	-10
	Semantic Condition: English Production											
300	667	690	-23	--	--	--	663	706	-43	706	723	-17
600	651	636	15	--	--	--	632	624	8	630	629	1

SOA [715 versus 753 ms; F (1, 14) = 12.6, p < .01]. Finally, there was a three-way interaction between the type of multiple translation, translation number, and SOA condition, F (1, 14) = 7.8, p< .05. This interaction shows that for words in the multiple form condition, there was a larger effect of SOA condition for the second translation than the first translation (101 versus 22 ms); for words in the multiple meaning condition there was a larger effect of SOA condition for the translation 1 condition than the translation 2 condition (58 versus 2 ms). Newman-Keuls tests showed that the form translation 2 and meaning translation 2 conditions showed significant relatedness effects.

For Dutch production, there was a significant effect of SOA condition such that words were named more quickly at the long SOA (703 ms) than at the short SOA [744 ms; $F(1, 14) = 16.1$, $p < .01$]. Two interactions were also significant. The interaction between type of multiple translation, translation number, and relatedness was significant, $F(1, 14) = 5.18$, $p < .05$]. This interaction showed that there were opposite effects of relatedness for words with multiple form and multiple meaning translations. For words with multiple form translations, there was a priming effect (related faster than unrelated by 21 ms) for the first translation and an inhibitory effect for the second translation (unrelated faster than related by 23 ms). For words with multiple meaning translations, there was an inhibitory effect for the first translations (unrelated faster than related by 24 ms), and a facilitatory effect for the second translations (related faster than unrelated by 34 ms). There was also an interaction between type of multiple translation, relatedness, and SOA condition, $F(1, 14) = 5.6$, $p < .05$. This interaction shows that for words with multiple form translations, the effect of SOA condition was larger for related (73 ms) than unrelated (34 ms) words, whereas for words with multiple meaning translations, there is a larger effect of SOA condition for unrelated (45 ms) than related (13 ms) words. Newman-Keuls tests showed that the meaning translation 2 and form translation 1 conditions showed relatedness effects. The meaning condition showed early facilitation and the form condition showed late facilitation. Because the same words are used in the two SOA conditions, these findings suggest that there may be phonological encoding of alternate forms/meanings in these conditions.

Note that the interactions with relatedness suggest that there may have been some phonological encoding during the Dutch production condition. The fact that the encoding is found in that condition and not in the English production condition may be due faster encoding of words in a bilingual's native language. However, the data from the translation trials (see below) suggest that if anything, translation from English to Dutch takes longer to perform than translation from Dutch to English, as was found in Experiment 1. Therefore, it seems unlikely that phonological encoding should occur at an earlier point in time during English to Dutch translation.

For the semantic conditions, naming trials from the meaning translation 2 condition were removed from this analysis because there was no corresponding form translation condition, and there was no significant difference between the two meaning translation conditions. This analysis therefore compared the form condition, the meaning translation 1 condition, and the one translation condition. There were three significant effects. First, there was an effect of relatedness, such that words were named more quickly in the related than unrelated conditions [647 versus 663 ms, F (1, 15) = 6.4, p < .05]. In addition, there was an effect of SOA such that words were named more quickly at the long than the short SOA [678 versus 632 ms, F (1, 15) = 43.9, p < .01]. Finally, there was a significant interaction between relatedness and SOA that qualifies these main effects [F (1, 15) = 4.9, p < .05]. Simple effects tests showed that the effect of relatedness was significant at the short SOA [F (1, 15) = 9.1, p < .01] but not the long SOA [F < 1, p > .10]. Thus, for the short SOA, there was an effect of relatedness, as expected, which is evidence for the activation of alternate conceptual candidates in this condition.

Translation Trials. The data for the translation trials were analyzed by SOA condition. It is not meaningful to compare the RTs from the form and meaning conditions because the words in these conditions were not matched in length and frequency. In the English production condition, translation was performed more slowly in the 300 ms condition than in the 600 ms condition [761 versus 635 ms; t (14) = 6.0, $p < .01$]. Similarly, in the Dutch production condition, translation was performed more slowly in the 300 ms condition than in the 600 ms condition [896 versus 704 ms; t (14) = 16.8, $p < .01$]. In the semantic condition, translation was performed more slowly in the 300 ms condition than in the 600 ms condition [719 versus 588 ms; t (15) = 8.8, $p < .01$]. It is to be expected that translation will be performed more quickly at the long SOA because the participants will have had more time to prepare their responses in that condition.

As in Experiment 1, translation from L2 to L1 was performed more slowly than translation from L1 to L2 (see Table 34). This finding again contrasts with the predictions of the revised hierarchical model. However, the finding of semantic relatedness effects in the L1 to L2 direction of translation coupled with the findings of other semantic effects in this direction of translation in Experiment 1, eliminates the possibility that the two directions of translation are qualitatively different and somehow the two languages switched dominance. If that were so, we would expect a reversed translation asymmetry, as found, however no semantic effects would be expected in L1 to L2 translation.

A model originally designed to explain task switching effects may be used to interpret the reversal of the translation asymmetry in this experiment (the Inhibitory Control Model; Green, 1998b). According to this model, translation from L1 to L2 is typically performed

Table 34

Translation Latencies as a Function of SOA and Condition

	SOA	
Condition	300 ms	600 ms
Phonological: English Production	761	635
Phonological: Dutch Production	896	704
Semantic: English Production	719	588

Note. There were no significant differences among the three groups in terms of length in Dutch, length in English, concreteness rating, cognate rating, or number of translations, *p*s > .10.

more slowly than translation from L2 to L1 because L1 to L2 translation requires the suppression of the L1 stimulus word. Because L1 presumably has more activation than L2, L1 will also require more suppression than L2. Interposing a production task during production in the dominant language (L2 to L1 translation) would be difficult to perform because the L1 word that had been prepared for production would be difficult to suppress. Therefore, the Inhibitory Control Model offers an alternative to the order effect explanation of the reversal of the translation asymmetry observed in this experiment.

Accuracy Analyses

The overall accuracy data from the naming and translation trials are given in Table 35. Translation was performed less accurately than naming, as would be expected. The only noticeable group difference is that the accuracy was highest in the semantic condition. The group of participants in that condition of the experiment also rated themselves as higher on L2 writing ability and speech comprehension ability than the participants in Experiment 1. It is possible that the overall accuracy trend is due to a difference in their ability or to the different stimuli used in the various conditions.

Table 35

Accuracy Data (%) for the Translation and Naming Trials of Experiment 3 by Condition

Condition	Mean Naming Accuracy	Mean Translation Accuracy
Phonological: English Production	92.0 (4.1)	71.9 (8.4)
Phonological: Dutch Production	91.6 (5.7)	72.1 (7.8)
Semantic: English Production	95.1 (2.4)	76.0 (6.7)

Note. Standard deviations are given in parentheses.

Effects of Number of Translations on Naming. The mean accuracy data for the critical trials by number of translations are given in Table 36. For production in English, two effects were significant. First, words were named more accurately at the long SOA than the short SOA [89 versus 95 %; F (1, 14) = 8.9, p < .01]. In addition, there was a three-way interaction between number of translations, type of multiple translation, and SOA, F (1, 14) = 5.0, p, .05. This interaction showed that the effect of SOA condition was not present for the related more than one translation condition (.4 %). For the other conditions, there was a larger effect that was similar for the unrelated more than one (9.2 %) and related one translation conditions (10.0 %), and somewhat smaller for the unrelated one translation condition (7.7 %).

For production in Dutch, there were three significant effects. First, words in the one translation condition were named less accurately than words in the more than one translation condition [89 versus 94 %; F (1, 14) = 11.3, p < .01]. Words were named more accurately at the short SOA than at the long SOA [93 versus 90 %; F (1, 14) = 6.8, p < .05]. Finally, there was an interaction between number of translations and SOA condition, F (1, 14) = 10.2, p <.01. Simple effects tests showed that the effect of SOA condition was significant for words with one translation [F (1, 14) = 11.5, p < .01], but not for words with more than one translation [F < 1, p > .10].

One effect was significant in the semantic condition. Words were named more accurately at the short than the long SOA [96 versus 94%; F (1, 15) = 6.2, p < .05].

Table 36

Accuracy Data (%) for Words with One and More than One Translation of Experiment 3 by Condition

	Number of Translations			
	One Translation		More than One Translation	
SOA	Related	Unrelated	Related	Unrelated
	Phonological Condition: English Production			
300 ms	96.7 (5.3)	96.1 (7.6)	91.5 (5.5)	97.0 (3.6)
600 ms	86.7 (16.6)	88.3 (12.9)	91.9 (5.5)	87.8 (9.0)
	Phonological Condition: Dutch Production			
300 ms	90.0 (12.6)	95.2 (5.5)	93.7 (8.1)	94.4 (5.6)
600 ms	87.0 (7.8)	85.6 (12.7)	92.2 (4.6)	94.4 (5.6)
	Semantic Condition: English Production			
300 ms	96.5 (4.9)	94.8 (4.3)	97.2 (3.5)	96.2 (4.4)
600 ms	94.1 (5.2)	94.1 (6.9)	94.4 (6.4)	93.1 (4.3)

Note. Standard deviations are given in parentheses.

Effects of Type of Multiple Translations on Naming. The mean accuracy data for the critical trials by type of multiple translations are given in Table 37. Two accuracy effects were found for production in English. First, there was a main effect of SOA such that words were named more accurately at the short SOA than at the long SOA [94 versus 89 %; F (1, 14) = 4.7, $p < .05$]. This effect is qualified by an interaction between type of multiple translation and SOA condition [F (1, 14) = 12.7, $p < .01$]. Simple effects tests showed that the effect of SOA was larger for unrelated words [10 %; F (1, 14) = 10.1, $p < .01$] than for related words [0 %; F (1, 14) < 1, $p > .10$].

Only one accuracy effect was significant for production in Dutch. This was the interaction of type of multiple translations and translation dominance, F (1, 14) = 5.3, $p < .05$. Simple effects tests showed that the effect of type was larger and in a different direction for the first translations [4 %; F (1, 14) = 4.0, $p = .06$] than for the second translations [-2 %; F (1, 14) = 1.6, $p > .10$]. There were no significant accuracy effects on naming in the semantic condition.

Table 37

Accuracy Data (%) for Words in the Multiple Form and Multiple Meanings Conditions of Experiment 3

	Type of Multiple Translations							
	Form				Meaning			
	Translation 1		Translation 2		Translation 1		Translation 2	
SOA	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
	Phonological Condition: Production in English							
300 ms	87.8	96.7	92.2	98.9	100.0	95.6	88.9	95.6
	(13.3)	(6.9)	(10.7)	(4.3)	(0)	(11.7)	(20.6)	(11.7)
600 ms	95.6	91.1	88.9	86.7	95.6	86.7	86.7	84.4
	(7.6)	(12.4)	(12.1)	(19.1)	(11.7)	(21.1)	(21.1)	(17.2)
	Phonological Condition: Production in Dutch							
300 ms	94.4	94.4	93.3	94.4	95.6	97.8	88.9	91.1
	(10.3)	(10.3)	(10.5)	(8.1)	(11.7)	(8.6)	(16.3)	(15.3)
600 ms	87.8	93.3	94.4	95.6	93.3	95.6	95.6	93.3
	(13.3)	(10.5)	(8.1)	(7.6)	(13.8)	(11.7)	(11.7)	(13.8)
	Semantic Condition: Production in English							
300 ms	97.4	94.8	--	--	97.9	97.9	95.8	100.0
	(4.0)	(6.0)			(8.3)	(8.3)	(11.4)	(0)
600 ms	96.4	92.2	--	--	93.8	93.8	87.5	95.8
	(6.1)	(6.4)			(13.4)	(13.4)	(16.7)	(11.4)

Note. Standard deviations are given in parentheses.

Translation Trials. The data for the translation trials were analyzed by SOA condition. In the English production condition, translation was performed more accurately in the 300 ms condition than in the 600 ms condition [73 versus 71 %; t (14) = 2.9, p < .05]. Similarly, in the Dutch production condition, translation was performed more accurately in the 300 ms condition than in the 600 ms condition [75 versus 69 %; t (14) = 3.1, p < .01]. In contrast, in the semantic condition, translation was performed more accurately in the 600 than the 300 ms condition [77 versus 75 %; t (15) = 2.1, p < .05].

Discussion

There are several notable findings of this experiment. First, the finding that translation was performed more quickly at the long than the short SOA suggests that participants did indeed begin to translate the words in this experiment because more preparation time shortened RTs. This finding is important because of the null effect of relatedness in the phonological conditions of this experiment for words with one translation. One could conclude that the participants did not begin to translate the words until the cue appeared, which would explain the lack of a relatedness effect for words with one translation. However, the SOA effect for translation along with the presence of semantic effects suggests that this is not the case. Instead, there may be some methodological explanation for the results, as will be discussed below.

The lack of a relatedness effect for words with one translation complicates the original interpretation of the data. When only one translation is available, all of the participants who were aware of the translation should have begun to translate the words in the same manner. This would have led to large relatedness effects based on phonological encoding. Although there may not be phonological encoding of any word other than the one the bilingual intends to say, the one translation condition should have shown this effect because there is only one option and because the participant must phonologically encode what he or she intends to say.

However, an alternative interpretation of this pattern of results is possible. In particular, because there are some conditions that demonstrate a relatedness effect, it is possible that the results for one translation words are a blend of facilitation and inhibition. For example, when there is a clear translation for the stimulus, the translation will be activated quickly, and the presentation of a naming target that is related to that stimulus will sometimes create a small amount of facilitation and sometimes create inhibition, perhaps depending on the stage of processing. However, when the naming target is related to the nondominant meaning/translation of the stimulus, and that nondominant meaning/translation is being suppressed, then naming will be delayed. The results from the phonological English production condition are consistent with this speculation, however additional research will be needed to determine whether this is a likely explanation for the results of this experiment.

In the English production condition, there was an effect of translation relatedness such that words in the translation 1 condition were named more quickly than words in the translation 2 condition. This finding suggests that despite the problems in the manner that translation number was assigned, the first translation was the dominant translation, at least on most trials. However, this distinction does not fit the hypotheses of this experiment because a difference between the dominant and subordinate translations was not anticipated in terms of reaction time. Instead, it was predicted that all translations would become equally activated (because the words had equi-probable translations). The data suggest the (slightly more) dominant translation is activated to a greater extent than the slightly less dominant translation. This difference in activation for the dominant and subordinate name is similar to what is found for within-language synonyms in picture naming (e.g., Peterson & Savoy, 1998). Again, the results from the phonological conditions must be interpreted with caution due to the lack of a relatedness effect for words with one translation.

In the Dutch production condition, there was an interaction between type of multiple translation, translation number, and relatedness that showed that words in the multiple meanings condition had inhibition for the first meaning and facilitation for the second meaning, whereas the reverse was true for words in the multiple forms condition. Furthermore, the interaction between type of multiple translation, relatedness, and SOA condition showed that the words with multiple form translations had a larger SOA effect for unrelated words than for related words, but that the reverse was true for words with multiple meaning translations. An examination of the difference scores shows that the first form translation leads to late facilitation (presumably because this word is then phonologically encoded). The second form translation leads to early inhibition. Perhaps this is because this form is activated only partially. The reverse pattern holds for the words with multiple meanings which show late inhibition for the first translation and early facilitation for the second translation. It is not clear why this would be the case.

The results of this experiment could be considered a failure to replicate the results of Peterson and Savoy (1998), however, there may be a methodological explanation for this. The null effect of relatedness in the phonological conditions suggests that the bilinguals in this case either did not begin performing translation, which seems unlikely. Alternatively, they may not have gotten to that point in the translation process, suggesting that the SOAs used were not in the proper range to observe these effects. It is most likely that the processes of interest occurred outside the range of SOA values (or between the two). Future research could address this by including a larger range of SOA values. Another possibility is that there was some flaw in the materials (e.g., that the related condition was not as related as intended) that led to the null effect.

The results of the semantic condition are the most clear to interpret. The findings show that words that are semantically related to the meanings of the stimulus are activated during translation. This finding supports the notion that translation involves conceptual mediation (e.g., De Groot, 1992). Furthermore, these results are consistent with those of Peterson and Savoy (1998) who reported that semantic associates of pictures are activated during picture naming, although these words do not become phonologically encoded. Finally, the presence of semantic effects suggests that the task used is an appropriate one but that the phonological effects may occur very late in the process and were not observed during the SOA values used in the present experiment.

Methodological Concerns

The major methodological concern with this experiment is that the SOA values that were selected were too short to observe the effects of interest in the phonological conditions of the experiment. Because semantic effects should occur earlier in the process of language production than the phonological effects, the fact that we find semantic effects in the absence of phonological effects is consistent with this conclusion. However, there was some evidence for phonological encoding of alternate translations that was not consistent across the conditions. Future research is needed to determine why these effects were not observed consistently.

Also, no attempt was made to match the probes in the various conditions in terms of whether they overlapped in terms of the onset or final syllable of the word (e.g., the difference between f*all*-t*all* and *err*and-*err*or). Because these two types of overlap have sometimes led to differential results (e.g., facilitation versus inhibition), the differences in the stimuli could explain why certain conditions showed facilitation instead of inhibition or vice versa.

Furthermore, the degree of overlap for probes in the various conditions should be measured subjectively and/or objectively (e.g., number of letters in same position) to ensure that probe relatedness does not differentially affect the conditions of interest. It is probably not possible to match words in length and frequency, as well as overlap across these conditions. Perhaps a solution would be to run several versions of the experiment using different sets of materials and seek convergence among them.

Finally, because the stimulus remained on the screen when the naming probe was presented, orthographic overlap may have influenced the results more than phonological overlap, and conscious checking processes could be implicated in the results. Thus, the results could be due to conscious pairing of the stimulus and probe and not to the automatic processing of the stimuli. In future research, a mask could be used to stop the processing of the first stimulus prior to the presentation of the second.

Conclusion

This experiment should be viewed as a first attempt to examine the time course of the number of translations effect. Overall, the results of this experiment did not show evidence for phonological encoding during the translation process, which was most likely due to the particular range of SOAs used. However, the activation of semantically-related concepts during the translation process was demonstrated, which is similar to the findings for picture naming within a language. The results from the phonological English production condition in particular may be taken to suggest that alternate translations may become activated during translation. To the extent that it is possible, future research should address the methodological issues listed above. If this can be done, the locus of the number of translations effect may be examined with fewer alternative interpretations of the data.

Chapter 5

General Discussion

Despite instances in which meaning clearly is not shared completely across languages (e.g., when languages have distinct lexical concepts), most psycholinguistic models of bilingual memory have ignored potential cross-language distinctions in meaning. Instead, these models assume that the two languages activate the same meanings, but to somewhat different degrees. This study was an attempt to further specify the nature of bilingual conceptual representation and to explore the consequences of imperfect meaning overlap across languages.

Because it is not possible to take all variables into account when doing research, it is often necessary to focus on only a limited range of phenomena, which results in many separate models that can each only account for a performance on a particular task or word type. Ultimately, however, a comprehensive model of cognition (or some subset of cognitive processing such as language processing) is desirable. To that end, this dissertation attempted to link issues in bilingual language processing with those in within-language processing. Although there are some models of language production and word recognition that focus on the ability of the bilingual to speak and comprehend two languages at different times, the present study was unique in that it focused on memory representation and the issues that may dictate meaning overlap across languages. Previous attempts to do this have been somewhat limited in the psycholinguistic domain.

Do Bilinguals Access the Same Meaning for Translation Equivalents?

The initial question of this thesis was whether bilinguals access the same meanings for words in two languages. The data from the present study did not directly address whether similar or different concepts are accessed by the two languages, but rather explored the consequences of different types of concepts for bilingual production. Although the results of the present study cannot ascertain whether there are similar concepts across languages, the data suggest that the meanings are *generally* quite similar for words in the two languages. First, bilinguals are able to translate words across languages, and this translation process appears to proceed via semantic access. This semantic access is evident in several semantic manipulations including the influence of conceptual salience as observed in Experiment 1 and the activation of semantic alternatives to the to-be-translated word that was observed in Experiment 3.

However, there is also evidence that suggests that bilinguals do not have *exactly* the same meanings for translation equivalents. First, bilinguals often transfer the meaning of an L1 word to its "translation equivalent" in L2 (e.g., Ijaz, 1986). In these cases, the bilinguals often ascribe meanings to the L2 word that are not appropriate because they derive the semantic boundary from their L1. For example, Ijaz showed that German speakers under-emphasized the contact component of the spatial preposition "on" and overemphasized the movement component. This problem presumably occurs because the close translation equivalent of "on" in German is "auf", which can denote a motional meaning like the English word “up”. Thus, these speakers have the appropriate central meaning of “on” but also an inappropriate non-central one.

Second, bilinguals sometimes do not transfer appropriate but non-central meanings of polysemous words from L1 to L2. Kellerman (1978) showed that Dutch-English bilinguals failed to transfer the intransitive meaning of “breken” to the English translation “break” despite its appropriateness in English. This finding shows that transfer of meaning is not complete in all cases, even when such transfer is appropriate.

What appears clear from these studies is that bilinguals sometimes do transfer meaning from L1 to L2, but that these meanings are primarily the core meanings of words. Perhaps more importantly, bilinguals often do not have the same semantic representations in their L2 as native speakers of the L2. A study by Zhang (1995) showed that ESL learners had different semantic structures from the semantic structures of native English speakers for degree adverbs (e.g., “extremely” to “slightly”) and frequency terms (e.g., “frequently if not always” to “once in a while”). However, increased proficiency in English led to closer approximation of native-like semantic structures. Thus, the semantics of the two languages seem to become more distinct with increased proficiency in the L2.

In sum, the available evidence suggests that on the whole, meanings are similar across languages, but that there may be language-specific features that form part of the meaning representations. Thus, there is a range of semantic similarity for words across languages. The point at which a translation pair falls along this continuum is likely to be due to a large set of factors, including but not necessarily limited to, conceptual salience, cognate status, and ambiguity (number of meanings and forms).

It is not known how meanings become differentiated across languages (i.e., more native-like in the L2) or what factors (linguistic or individual) may facilitate or discourage such differentiation. Furthermore, most studies have focused on the distinctions for the L2 being acquired, but have not examined the potential consequences of semantic differentiation for general conceptual representation. If a bilingual truly does restructure a meaning to correspond more closely to the L2 structure, either the meaning is specific to the L2 or there is some permanent change that pervades meaning representation in general. For example, when a native English speaker learns the two verbs for "to be" in Spanish, does she or he restructure the verb into two components (e.g., a more temporary state to correspond to "estar" and a more permanent state to correspond to "ser")?

Furthermore, when bilinguals become more native-like in their use of L2 vocabulary/syntactic structures, have their intuitions about the language changed (or have the semantics been restructured), or are they better able to apply rules to language use? A study of the electrophysiological responses to improper forms during comprehension could help determine the answer to these questions by revealing whether bilinguals are sensitive to the incorrectness of the wrong form during initial comprehension.

Modeling Culturally-Distinct Meanings Across Languages

How might linguistic/cultural distinctions in meaning be expressed in a model of bilingual memory? The data presented here suggest that there may be word-specific factors that determine the relative overlap of meanings across languages (e.g., number of meanings, number of synonyms, conceptual salience, cognate status). Thus, as De Groot (1992) proposed, the overlap of translation equivalents at the semantic level may depend on word-specific characteristics such as conceptual salience. However, on the basis of the present study, it appears that the view adopted in De Groot's work is somewhat too simplistic. This is because concreteness (or conceptual salience) and cognate status are not sufficient to define the full range of semantic distinctions across languages. In addition to these factors, number of synonyms and number of meanings must be considered, especially because they interact with conceptual salience and cognate status. Other factors such as conceptual animacy may also need to be taken into account. Furthermore, as noted by Paradis (1997), differences can be found in the referents of translation equivalents of concrete words (e.g., the English "ball" and the French "balle").

Despite the potential problems with the particular instantiation of the distributed feature model (De Groot, 1992), the idea behind that model is potentially very useful. Thus, it may not be that cognates and concrete words are translated more quickly than noncognates and abstract words because the former share more conceptual features across languages than the latter. Rather, other factors may be responsible for these patterns of results. In the case of concrete words, concrete word meanings are accessed more readily within a language, and therefore there is no need to postulate an additional cross-language mechanism to explain concreteness effects in translation. In the case of cognate translations, there is orthographic and phonological similarity of the translation equivalents that could also be responsible for faster translation, at least partially. Most convincingly, the results of the present study show that, in some cases, words low in conceptual salience are translated as quickly as (or more quickly than) words high in conceptual salience. This elimination or reversal of the typical concreteness effect suggests that the simple assumption that concrete words are more similar than abstract words across languages is not sufficient to explain translation performance.

Rather than postulating that different types of concepts are more or less likely to share features across languages, particular concepts should ideally be investigated on an individual basis. The available evidence suggests that other important factors will need to be taken into account in determining the likelihood that words will share conceptual features with their translation equivalents across languages. In particular, one important factor will be the semantic range of the word's meaning (e.g., Kellerman, 1978). Based on De Groot's distributed feature model, it can be speculated that there are sets of distributed features that comprise meaning, and that some features are uniquely connected to a word (or words) in one of the bilingual's languages.

Other suggestions that have been made in the literature on conceptual representation in general indicate other possibilities for modeling cultural distinctions in meaning. Barsalou (1982) proposed that words have context-independent and context-dependent properties or features. Context-independent features of words are those that become activated whenever the word is encountered. In contrast, the context-dependent properties of words only become activated in an appropriate context. For example, the word "ball" has the properties "is round" and "floats". The roundness feature would presumably be activated regardless of the context in which "ball" appears because it is a defining feature of "ball"; however, the float feature would only be relevant in particular contexts such as when the ball falls into water. Thus, Barsalou's hypothesis suggests that some properties of words are core or central versus other properties that are peripheral or do not define the meaning of the word. Perhaps the core meaning features (i.e., context-independent) are shared across languages, but the more peripheral ones (i.e., context-dependent) are not. If this were true, translation equivalents would have enough shared features to be called translations at a very basic level, but the culturally-specific contexts in which these words appear and the unique features associated with the word in each language would be preserved.

Barsalou's proposal could also be used to explain a way that the conceptual representations may change as a function of acquiring an L2. Instead of restructuring the existing concepts entirely (e.g., when a native English speaker learns the two verbs "to be" in Spanish), perhaps the change is in the form of minor elaborations (context-dependent) to the core (context-independent) concepts.

Similarly, McRae, De Sa, and Seidenberg (1997) and McRae and Boisvert (1998) proposed that concrete words have perceptual features as part of their meanings (e.g., the features 'has a tail' and 'has fur' for the concept "dog"). Furthermore, at least for living things, such perceptual features tend to be correlated. For example, many animals that have tails also have fur and four paws. McRae et al. demonstrated the importance of perceptual features for word meanings and the effects on translation were demonstrated by Sholl (1995) who showed that animate concepts were translated more quickly than inanimate concepts. It is possible that the perceptual features of words (correlated or not) are shared across languages, but other features (e.g., that indicate the contexts in which objects are typically used) are uniquely associated with one language.

A final suggestion is that there are a set of conceptual primitives upon which one draws to infer meaning. There would be one set of primitives for the two languages, but these primitives would combine somewhat differently for words in the two languages. This would result in subtle meaning differences for translation equivalents across languages (in certain cases), but would still maintain the core meanings of words across languages. This proposal is similar to the one put forth by Barsalou (1982), however it differs in that no specific features or primitives would necessarily be accessed upon *every* occurrence of the word.

Whatever the particular form of the semantic representation, it is too simplistic to assume that entire categories of concepts have translation equivalents that are more or less similar across languages. A more specific analysis of words in various cultures would be most beneficial to the study of these ideas. Such an analysis could include particular translation equivalents that seem to be culturally-specific. However, it is quite likely that the specificity is dependent upon the language history of the particular bilinguals; the way in which a person acquires a second language will influence their awareness of and sensitivity to cultural nuances in meaning. Learners who are immersed in a second language culture and/or learners who are in communication with individuals from that culture will be more aware of differences than, for example, classroom learners who are not instructed on cross-cultural differences (although some clearly are).

Implications for Models of Cross-Language Processing

The results of Experiment 1 showed that conceptual salience, cognate status, and cross-language ambiguity influence translation performance. In general, translation pairs high in their orthographic/phonological overlap were translated more quickly than translations low in overlap. Furthermore, words high in conceptual salience were translated more quickly than words low in conceptual salience. In addition, words with multiple meanings or multiple synonyms were translated more slowly than unambiguous words. These main effects were qualified by an interaction between conceptual salience, cognate rating, and ambiguity. The nature of this interaction indicates that multiple forms slow translation regardless of conceptual salience and cognate rating. In contrast, for noncognates, multiple meanings slow translation only for words high in conceptual salience, whereas for cognates, multiple meanings slow translation regardless of conceptual salience. A model of conceptual salience and ambiguity effects was presented that accounted for the results in terms of a language processing system in which meaning activation is a function of conceptual salience and all forms for a word become activated during the translation process.

This model makes certain assumptions regarding the effects of conceptual salience in the bilingual translation context. In particular, the meanings of words high in conceptual salience will be activated more quickly than the meanings of words low in conceptual salience. However, the model does not make any assumptions about particular word types being more or less similar across languages.

These results have implications for current theories of conceptual salience effects because the findings show that conceptual salience effects in translation are not ubiquitous, but rather depend on the ambiguity of the stimuli. Furthermore, conceptual salience effects have been used as an index of meaning activation during various cross-language tasks (e.g., translation production, translation recognition, word association). These effects play a large role in the debate regarding the degree to which the two languages/directions of translation invoke meaning because they have been one of the main sources of evidence used to suggest that both directions of translation are conceptually mediated. However, because conceptual salience effects depend on the ambiguity of the stimuli, these effects cannot be interpreted without further qualification of the conditions under which these effects are observed.

Implications for the Study of Bilingual Language Processing

The results of this study indicate consequences of certain types of concepts for bilingual production. In particular, it is evident that cross-language ambiguity must be taken into account in considering bilingual language processing, and most likely language learning. Furthermore, the research in the area of bilingual representation has typically focused on studies including words with pictureable referents. This area of research needs to be extended to new domains of stimuli as many interesting differences can only be observed when a range of word types are used.

Experiment 3 was performed to examine the time course of the ambiguity effects observed in Experiment 1. Although the results of this experiment were somewhat difficult to interpret, the overall pattern of results suggests that the methodology employed may allow for the examination of the time course of translation. It will be critical to extend the SOA values to cover a wider range of time, and also to ensure that there are no stimulus characteristics that may reduce or eliminate the effects of interest (e.g., the type of overlap between the related words). It can be concluded on the basis of these data that semantic associates of words are activated during translation as in within-language picture naming (e.g., Peterson & Savoy, 1998). It can further be speculated that multiple translations are phonologically encoded during translation due to the relatedness effects observed in the phonological English production condition. In the future, the interaction of context and these effects should be examined so that the effects can be placed in a more realistic language processing setting.

Implications for Models of Within-Language Processing

The results of Experiment 1 showed that the effects of conceptual salience were modulated by ambiguity. It was proposed that if these effects were a general property of the language processing system, that similar effects would be found in a within-language task. The results of the present Experiment 2 suggest that lexical access is influenced by both conceptual salience and number of meanings, however no interaction between the factors was observed. One interpretation of the discrepant results between the two experiments is that the results of Experiment 1 were specific to bilingual language processing. However, a somewhat more likely explanation is that the differences in the task demands between the two experiments led to the different patterns of results. This possibility can be examined in future research by using a within-language task that requires production (e.g., animacy effects in picture naming) or by comparing performance on the within and cross-language lexical decision tasks.

This study is the first to examine the implications of conceptual salience and ambiguity for lexical access concurrently. Past research may have led to the conclusion that number of meanings does not predict performance because of the confounding of conceptual salience and ambiguity. However, the results of Experiment 2 suggest that both conceptual salience and ambiguity are important to within-language processing, and that these factors operate independently. Because past studies did not take both conceptual salience and number of meanings into account simultaneously, the results that did not show NOM effects may have been due to the confounding of conceptual salience and ambiguity which masked the ambiguity effects. It is not possible to determine this without obtaining conceptual salience ratings on the words used in these past studies, and stimulus lists are not available for many studies.

The results of Experiment 2 also make evident the importance of task demands on performance. Effects of task demands place certain constraints on the generalizability of research findings, however they also demonstrate the flexibility of the human language system to adapt to the task at hand. Future research in this area should consider the precise effects of context on the types of effects observed in this experiment because previous studies have not considered ambiguity and conceptual salience simultaneously.

Future Directions

In future research, it will be important for the ambiguity of the stimuli to be considered, especially in terms of the translation task. This variable is just one that shows the importance of the cross-language overlap. Future models of bilingual language processing will need to take these issues into account as they clearly determine translation performance. The importance of stimulus ambiguity for language learning should also be explored further.

Furthermore, context effects should be examined in the cross-language and within-language domains. It will be important to understand how context aids in the disambiguation of ambiguous words and also how context aids in the comprehension of words low in conceptual salience.

Finally, the time course of translation should be studied more extensively and with a wider range of SOAs than were used in the present study. The paradigm used in Experiment 3 has the potential to show the effects of ambiguity and conceptual salience as they unfold over time, but the methodological concerns raised earlier will need to be addressed.

References

Aiken, , L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage Publications, Inc.

Altarriba, J. (1990). *Constraints on interlingual facilitation effects in priming in Spanish-English bilinguals.* Unpublished dissertation, Vanderbilt University.

Altarriba, J., & Mathis, K. M. (1997). Conceptual and lexical development in second language acquisition. *Journal of Memory and Language, 36*, 550-568.

The American heritage electronic dictionary (3rd ed.) [Computer software]. (1992). Houghton Mifflin Co.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6,* 451-474.

Anderson, J. R. (1983) Retrieval of information from long-term memory. *Science, 220,* 25 30.

Azuma, T. & Van Orden, G. C. (1997). Why SAFE is better than FAST: The relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language, 36,* 484-504.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database (Release 2) [CD-ROM]*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10,* 82-93.

Begg, I., & Paivio, A. (1969). Concreteness and imagery in sentence meaning. *Journal of Verbal Learning and Verbal Behavior, 8,* 821-827.

Chen, H.-C., & Ng, M.-L. (1989). Semantic facilitation and translation priming effects in Chinese-English bilinguals. *Memory & Cognition, 17,* 454-462.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25,* 257-271.

Costa, A., Miozzo, M., Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41,* 365-397.

De Groot, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1001-1018.

De Groot, A. M. B. (1993). Word-type effects in bilingual processing tasks: Support for a mixed-representational system. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon: Vol. 6. Studies in bilingualism* (pp. 27-51). Amsterdam, Netherlands: John Benjamins Publishing Company.

De Groot, A. M. B. (1995). Determinants of bilingual lexicosemantic organization. *Computer Assisted Language Learning, 8,* 151-180.

De Groot, A. M. B. Dannenburg, L., & Van Hell, J. G (1994). Forward and backward word translation by bilinguals. *Journal of Memory and Language, 33,* 600-629.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30,* 90-123.

De Groot, A. M. B., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning, 47*, 215-264.

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review, 98,* 604-614.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41,* 496-518.

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism, 1,* 51-66.

Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition*, *23*, 166-180.

Dyer, F. N. (1971). Color-naming interference in monolinguals and bilinguals. *Journal of Verbal Learning and Verbal Behavior, 10,* 297-302.

Ellis, H. D., & Shepherd, J. W. (1974). Recognition of abstract and concrete words presented in left and right visual fields. *Journal of Experimental Psychology, 103,* 1035-1036.

Elsinger, C. L. (1996). *Accessing concepts for second language words.* Unpublished master's thesis, The Pennsylvania State University, University Park, PA.

Forster, K. I., & Bednall, E. S. (1976). Terminating and exhaustive search in lexical access. *Memory & Cognition, 4,* 53-61.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627-635.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin Co.

Francis, W. S. (1999a). Analogical transfer of problem solutions within and between languages in Spanish-English bilinguals. *Journal of Memory and Language, 40,* 301-329.

Francis, W. S. (1999b). Cognitive integration of language and memory in bilinguals: Semantic representation. *Psychological Bulletin, 125,* 193-222.

Frenck, C., & Pynte, J. (1987). Semantic representation and surface forms: A look at across-language priming in bilinguals. *Journal of Psycholinguistic Research, 16,* 383-396.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113,* 256-281.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1122-1139.

Grabois, H. (1999). The convergence of sociocultural theory and cognitive linguistics: Lexical semantics and the L2 acquisition of love, fear, and happiness. In G. B. Palmer & D. J. Occhi (Eds.), *Languages of sentiment: Cultural constructions of emotional substrates (pp. 201-233)*. Amsterdam, Netherlands: John Benjamins Publishing Co.

Green, D. W. (1998a). Bilingualism and thought. *Psychologica Belgica, 38.*

Green, D. W. (1998b). Mental control of the bilingual lexico-semantic system. *Bilingualisn Language and Cognition, 1*, 67-81.

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development, 6,* 467-477.

Grosjean, F. (1997). Processing mixed language: Issues, findings, and models. In A. M. B. De Groot and J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 225-254). Hillsdale, NJ: Lawrence Erlbaum Publishers.

Hancin-Bhatt, B., & Nagy, W. (1994). Lexical transfer and second language morphological development. *Applied Psycholinguistics, 15,* 289-310.

Heredia, R. R. (1997). Bilingual memory and hierarchical models: A case for language dominance. *Current Directions in Psychological Science, 6,* 34-39.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism, 1,* 213-229.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1331-1356.

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed representations. In D. E. Rumelhart, J. E. McClelland, and the PDP Research Group (Eds.), *Parallel distributed processing. Explorations in the Microstructure of Cognition, Volume 1: Foundations* (pp. 77-109). Cambridge: MIT Press.

Holmes, V. M., & Langford, J. (1976). Comprehension and recall of abstract and concrete sentences. *Journal of Verbal Learning and Verbal Behavior, 15,* 559-566.

Ijaz, H. (1986). Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning, 36,* 401-451.

Jacobs, A. M., & Grainger, J. (1992). Testing a semistochastic variant of the Interactive Activation model in different word recognition experiments. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1174-1188.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 104,* 130-136.

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology, 13,* 278-305.

Jastrzembski, J., & Stanners, R. (1975). Multiple word meanings and lexical search speed. *Journal of Verbal Learning and Verbal Behavior, 14,* 534-537.

Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: Evidence from the coactivation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1256-1274.

Jiang, N. (1999). Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism, 2,* 59-75.

Keatley, C., Spinks, J., & De Gelder, B. (1994). Asymmetrical semantic facilitation between languages. *Memory & Cognition, 22,* 70-84.

Kellerman, E. (1978). Giving learners a break: Native language intuitions as a source of predictions about transferability. *Working Papers on Bilingualism, 15,* 59-92.

Kieras, D. (1978). Beyond pictures and words: Alternative information-processing models for imagery effects in verbal memory. *Psychological Bulletin, 85,* 532-554.

Kirsner, K., Smith, M. C., Lockhart, R. S., King, M. L., & Jain, M. (1984). The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior, 23,* 519-539.

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 92-107.

Kroll, J. F., & Stewart, E. (1994). Concept interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149-174.

La Heij, W., De Bruyn, E., Elens, E., Hartsuiker, R., Helaha, D., & Van Schelven, L. (1990). Orthographic facilitation and categorical interference in a word-translation variant of the Stroop task. *Canadian Journal of Psychology, 44,* 76-83.

La Heij, W., Kerling, R., & Van der Velden, E. (1996). Nonverbal context effects in forwar and backward translation: Evidence for concept mediation. *Journal of Memory and Language, 35,* 648-665.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98,* 122-142.

Lucy, J. A. (1997). Linguistic relativity. *Annual review of anthropology, 26,* 291-312.

MacIntyre, P. D., Noels, K. A., & Clement, R. (1997). Biases in self-ratings of second language proficiency: The role of language test performance. *Language Learning, 47,* 265-287.

MacWhinney, B. (1997). Second language acquisition and the competition model. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives (pp. 113-142)*. Hillsdale, NJ: Lawrence Erlbaum Publishers.

MacWhinney, B. (in press). The competition model: The input, the context, and the brain. Chapter to appear in P. Robinson (Ed.) *Cognition and second language instruction.* Cambridge University Press.

Mägiste, E. (1984). Stroop tasks and dichotic translation: The development of interference patterns in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 304-315.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375-407.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114,* 159-188.

McElree, B., Jia, G., & Litvak A. (2000). The time course of conceptual processing in three bilingual populations. *Journal of Memory and Language, 42,* 229-254.

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24,* 558-572.

McRae, K., De Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126,* 99-130.

Meyer, D. E., & Ruddy, M. G. (1974, April). *Bilingual word recognition: Organization and retrieval of alternative lexical codes.* Paper presented at the 45th Annual Meeting of the Eastern Psychological Association, Philadelphia, PA.

Michael, E. B. (1998). *The consequences of individual differences in cognitive abilities for bilingual language processing.* Unpublished dissertation, The Pennsylvania State University, University Park, PA.

Miller, N. A. (1997). *The influence of word form and meaning in bilingual language production.* Unpublished master's thesis, The Pennsylvania State University, University Park, PA.

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory & Cognition, 17,* 141-147.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition, 9,* 225-236.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart, & Winston.

Paradis (1997). The cognitive neuropsychology of bilingualism. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives (pp. 331-354)*. Hillsdale, NJ: Lawrence Erlbaum Publishers.

Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism, 2,* 209-230.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 539-557.

Piercey, C. D., & Joordens, S. (2000). Turning an advantage into a disadvantage: Ambiguity effects in lexical decision versus reading tasks. *Memory & Cognition, 28,* 657-666.

Potter, M. C., So, K-F., Von Eckhardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23,* 23-38.

Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color-word task. *Journal of Verbal Learning and Verbal Behavior, 8,* 295-301.

Prisma Woordenboek: Engels-Nederlands (1990). Utrecht, The Netherlands: Uitgeverij Het Spectrum.

Prisma Woordenboek: Nederlands-Engels (1990). Utrecht, The Netherlands: Uitgeverij Het Spectrum.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences, 24,* 574-590.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510-532.

Rubenstein, H., Garfield, L., & Millikan, J. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9,* 487-494.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10,* 645-657.

Sánchez-Casas, R. M., Davis, C. W., & García-Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate/non-cognate distinction. *European Journal of Cognitive Psychology, 4*, 293-310.

Schönpflug, U. (1997, April). *Bilingualism and memory.* Paper presented at the International Symposium on Bilingualism, Newcastle-upon-Tyne, United Kingdom.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language, 27,* 499-520.

Schwanenflugel, P. J., & Rey, M. (1986). Interlingual semantic facilitation: Evidence for a common representational system in the bilingual. *Journal of Memory and Language, 25,* 605-618.

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 82-102.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14,* 489-537.

Sholl, A. (1995). *Animacy effects in picture naming and bilingual translation: Perceptual and semantic contributions to concept-mediation.* Unpublished dissertation, The University of Massachusetts, Amherst.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer between picture naming and translation: A test of asymmetries in bilingual memory. *Psychological Science, 6*, 45-49.

Sinclair, J. M. (Ed.). (1987). *Looking Up.* London/Glasgow: Collins/COBUILD.

Spreen, O., & Schulz, R. W. (1966). Parameters of abstraction, meaningfulness, and pronunciability for 329 nouns. *Journal of Verbal Learning and Verbal Behavior, 5,* 459-468.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18,* 645-659.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language, 27,* 324-340.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior, 18,* 427-440.

Tokowicz, N., & Kroll, J. F. (2000). *Accessing meaning for words in two languages: The effects of concreteness and multiple translations in bilingual production.* Manuscript submitted for publication.

Tzelgov, J., & Eben-Ezra, S. (1992). Components of the between-language semantic priming effect. *European Journal of Cognitive Psychology, 4,* 253-272.

Tzelgov, J., & Henik, A. (1990). The insensitivity of the semantic relatedness effect to surface differences and its implications. In P. J. Drenth & J. A. Sergeant (Eds.), *European perspectives in psychology, Vol. 1: Theoretical, psychometrics, personality, developmental, educational, cognitive, gerontological* (pp. 385-402). Chichester, England: John Wiley & Sons.

Van Hell, J. G. (1998). *Cross-language processing and bilingual memory organization.* Unpublished dissertation, University of Amsterdam.

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

Van Hell, J. G., & De Groot, A. M. B. (1998b). Disentangling context availability and concreteness in lexical decision and word translation. *Quarterly Journal of Experimental Psychology (A), 51,* 41-63.

Van Hell, J. G., & Dijkstra, T. (2000). *Foreign language knowledge can influence native language performance: Evidence from trilinguals.* Manuscript submitted for publication.

Whorf, B. L. (1956). Science and linguistics. In J. B. Carroll (Ed.), *Language, thought and reality: Selected writings of Benjamin Lee Whorf.* Cambridge, MA: MIT Press.

Zhang, S. Q. (1995). Semantic differentiation in the acquisition of English as a second language. *Language Learning, 45,* 225-249.

Appendix A

Language History Questionnaire

Sex: M / F Age (in years) ______ Native country ______

Years spent in the U.S. ______ Years spent in U.S. schools ______

This questionnaire is designed to give us a better understanding of your experience learning a second language. We ask that you be as accurate and thorough as possible when answering the following questions and thank you for your participation in this study.

If at any time, you need more space to write, please feel free to use an extra sheet of paper. Please put the question number beside your responses.

1.) Do you have any known visual or hearing problems (corrected *or* uncorrected)?

2.) What is your first language (i.e., language first spoken)? If more than one, please briefly describe the situations in which each language was used.

Which language do you consider your *second* language? _____________

3.) If you have ever lived in or visited a country where languages other than your native language are spoken, please indicate below the name of the country (countries), the duration of your stay in number of months, and which languages you used while you were in the country (please indicate if you were spoken to in a language other than your first language, even if you never actually spoke that language).

Country visited	*# Months there*	*Language(s) used*

4.) List below, from most fluent to least fluent, *all* of the languages you know. Also specify the age in years at which you began to learn the language and the context in which you learned it. For example, "Dutch, birth, home". Include *all* languages to which you have been exposed, although you may never have had formal training in them and may not be able to read, speak or write them.

Please remember to list your native language.

Language	*Age in yrs.*	*Learning Situation*
Dutch		
English		

5.) What languages were spoken in your home while you were a child and by whom?

6.) How many years have you studied English? Please indicate the setting(s) in which you have had experience with the language (i.e., classroom, with friends, foreign country...)

Number of years:

Setting(s):

7.) What languages other than Dutch do you *speak* fluently?

8.) What languages other than Dutch do you *read* fluently?

9.) What languages other than Dutch do you *write* fluently?

10.) What languages other than Dutch do you *understand* when they're spoken?

11.) Are you currently enrolled in any language courses? (Any course either instructed in a foreign language or designed to teach a foreign language). If so, please list the course number(s) below, along with the title of the course.

Department	*Course #*	*Title of Course*

12.) Please list the language(s) the following people speak.

Mother:

Father:

Closest friend:

For the next eight questions, please circle the number of your response:

13.) Please rate your English *reading* proficiency on a ten-point scale.
(1= not literate, 10= very literate)

1 2 3 4 5 6 7 8 9 10

not literate very literate

14.) Please rate your Dutch *reading* proficiency on a ten-point scale.
(1= not literate, 10= very literate)

1 2 3 4 5 6 7 8 9 10

not literate very literate

15.) Please rate your English *writing* proficiency on a ten-point scale.
(1= not literate, 10= very literate)

1 2 3 4 5 6 7 8 9 10
not literate very literate

16.) Please rate your Dutch *writing* proficiency on a ten-point scale.
(1= not literate, 10= very literate)

1 2 3 4 5 6 7 8 9 10
not literate very literate

17.) Please rate your English *conversational fluency* on a ten-point scale.
(1= not fluent, 10= very fluent)

1 2 3 4 5 6 7 8 9 10
not fluent very fluent

18.) Please rate your Dutch *conversational fluency* on a ten-point scale.
(1= not fluent, 10= very fluent)

1 2 3 4 5 6 7 8 9 10
not fluent very fluent

19.) Please rate your English *speech comprehension ability* on a ten-point scale.
(1= unable to understand conversation, 10= perfectly able to understand conversation)

1 2 3 4 5 6 7 8 9 10
no comprehension perfect comprehension

20.) Please rate your Dutch *speech comprehension ability* on a ten-point scale.
(1= unable to understand conversation, 10= perfectly able to understand conversation)

1 2 3 4 5 6 7 8 9 10
no comprehension perfect comprehension

21.) How do you consider your foreign language learning skills? Please circle your response.

worse than average average better than average

22.) When learning a foreign language, which of the following do you find the easiest to learn? Please rank the following from 1 to 4. (1= easiest, 4 = hardest)

Pronunciation: ______

Grammar: ______

Vocabulary: ______

Sayings/Expressions: ______

23.) Have you ever lived or stayed somewhere for a month or more where English was spoken as a first language (e.g., England, the United States)?
(please circle one) yes no

If yes, please answer questions 24 - 36 below. If no, please skip to question 37.

24.) What was your experience with English prior to staying in a country where it was spoken? (for example, how many courses had you taken, what level courses had you taken, had you ever spoken the language with a native speaker of that language...)

____
____
____
____

** While you were in the English speaking environment: **

25.) Did you attend any structured course(s)? If so, please list them below. For each course, please list the primary language of instruction (if more than one was used, please list *all* languages used.)

____
____
____
____

26.) How many hours did you spend in each class *each day*? What was the total number of hours spent in each class *per week*?

____
____

27.) Do you feel that being in an English-speaking environment has had an effect on your *first* language (Dutch)? Please explain.

____
____
____
____

28.) Were you actively encouraged to suppress your first language (Dutch) while abroad? If so, by whom?

29.) Please describe, in detail, your living situation while abroad. For example, did you live with a family or in a dorm? If you did not live alone, which language(s) did you speak with the people you lived with? Were you traveling as on vacation?

30.) While in the immersion environment, did you speak English with other non-native speakers? If so, what were the circumstances? (i.e., did you not share the same first language?, others were present who did not understand your first language?, etc.)

31.) While abroad, did you ever think or dream in English? If so, was this the first time that this had ever happened? Has it happened since your return? Please explain.

32.) While you were in the English-speaking environment,
What percentage of the time did you spend speaking English ? ____%
What percentage of the time did you spend speaking Dutch? ____%

33.) How much time did you spend watching TV in English per day (please circle)?
less than 1 hour 1-2 hours 3-4 hours more than 4 hours

34.) How frequently did you read a newspaper or magazine in English (please circle)?
never once a week three or more times a week

35.) How much time did you spend speaking with native speakers of English?
hardly ever occasionally often at every opportunity

36.) Please compare learning a second language in an immersion (e.g., English-speaking environment with learning one in a classroom environment. (Which is easier? In which did you learn more?...) Please comment about the difference between your various learning experiences.

37.) Is there anything else about your language background that you would like to comment on? Please feel free to make comments about things which were not covered on this questionnaire.

THANK YOU FOR YOUR PARTICIPATION !

Appendix B

Language History Questionnaire Data from the Participants in all Bilingual Tasks

	Task					
Measure	Number of Translations Norms	Cognate Norms	Exp. 1 Translation	Exp. 3 English Production	Exp. 3 Dutch Production	Exp. 3 Semantic Condition
Age (years)	21.6 (3.0)	21.2 (2.3)	21.6 (2.1)	22.5 (3.8)	22.3 (2.9)	21.9 (2.0)
Age Began L2 (years)	10.5 (2.1)	9.0 (2.5)	9.9 (3.2)	8.7 (3.7)	10.3 (2.0)	9.3 (2.6)
Time Studied L2 (years)	9.3 (3.6)	9.9 (3.8)	9.7 (3.7)	11.0 (3.8)	9.9 (3.0)	11.4 (2.6)
L2 Immersion Experience (months)	5.0 (12.3)	2.0 (4.0)	2.3 (4.3)	1.9 (4.7)	2.5 (4.5)	3.6 (6.6)
L2 Reading Ability	7.5 (1.3)	8.2 (.9)	8.4 (.9)	8.4 (1.1)	8.1 (1.2)	8.7 (.7)
L2 Writing Ability	6.5 (1.3)	7.1 (1.1)	6.9 (1.7)	7.5 (1.3)	7.1 (1.2)	7.6 (.6)
L2 Conversation Ability	6.9 (1.2)	7.3 (1.3)	7.4 (1.6)	7.7 (1.2)	7.7 (1.3)	7.6 (1.1)
L2 Speech Comprehension Ability	8.0 (1.0)	8.2 (.9)	8.8 (1.1)	8.7 (.9)	8.4 (1.2)	9.0 (.5)

Note. Standard deviations are given in parentheses. There were no significant differences among the groups in terms of age, age began L2, time studied L2, L2 immersion experience, L2 writing ability, and L2 conversational ability, *p*s > .10. There was a significant difference among the groups in terms of L2 writing ability [F (5, 103) = 2.3, $p < .05$]. According to Newman-Keuls tests, the participants in the number of translations norming task were significantly lower than those in the Experiment 3 semantic condition at the $p = .05$ level. Furthermore, there was a significant difference among the groups in terms of L2 speech comprehension ability [F (5, 103) = 2.5, $p < .05$]. Again, Newman-Keuls tests showed that the participants in the number of translations condition were significantly lower than those in the Experiment 3 semantic condition at the $p = .05$ level.

Appendix C

Instructions for Cognate Ratings (Adapted from De Groot & Nas, 1991)

Many words in Dutch and English share sound and/or spelling in the two languages. The following word pairs consist of a Dutch word and one of its English translations.

Your first task is to rate the similarity of the two words in terms of their meaning. The rating scale goes from 1, which indicates “completely different” to 7, which indicates “exactly the same”.

Your second task is to rate the similarity of each word pair in terms of spelling and sound. The rating scale goes from 1, which indicates “low similarity” to 7, which indicates “high similarity”. *Your rating should reflect a combination of both the spelling and sound similarity.*

Examples:

		Meaning							Spelling/Sound						
		completely different				exactly the same			low similarity				high similarity		
jurk	dress	1	2	3	4	5	6	**7**	**1**	2	3	4	5	6	7
pen	pen	1	2	3	4	5	6	**7**	1	2	3	4	5	6	**7**
arrival	citroen	**1**	2	3	4	5	6	7	**1**	2	3	4	5	6	7

If you are not sure how to rate a word pair, it is appropriate to guess or follow your first instinct. Please rate the items in the order in which they appear in the list. Rate each pair in terms of meaning first, then spelling/sound. Please do not change your responses or go back to a previous item.

Appendix D
Scatterplots of the Predicted Values against Residuals from Experiment 1

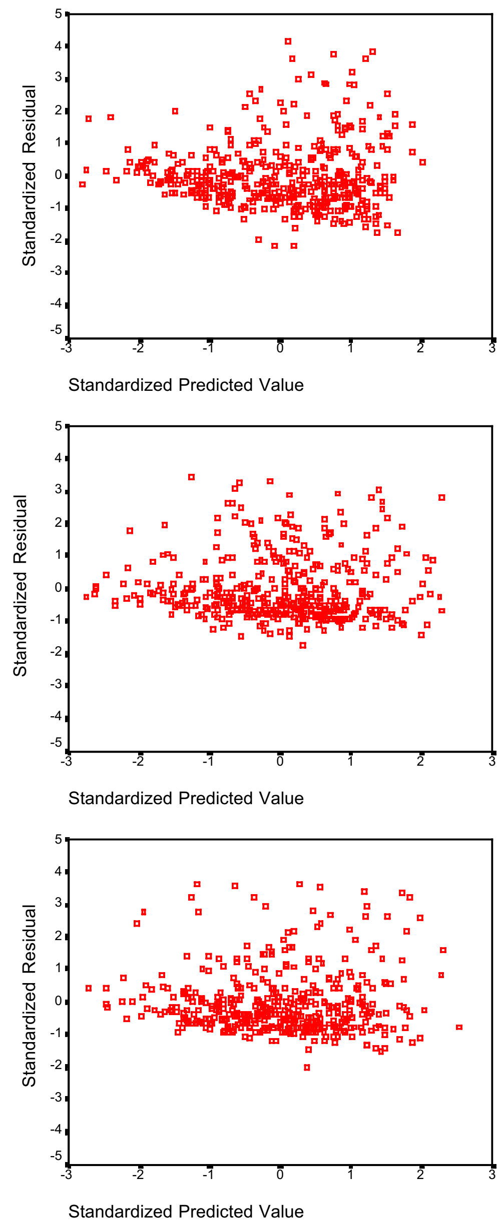


Figure 12. Scatterplot of predicted values against residuals from the direction of translation analysis (taken from participants 1, 2, and 3 in the top, middle, and bottom panels, respectively).

Appendix D, continued

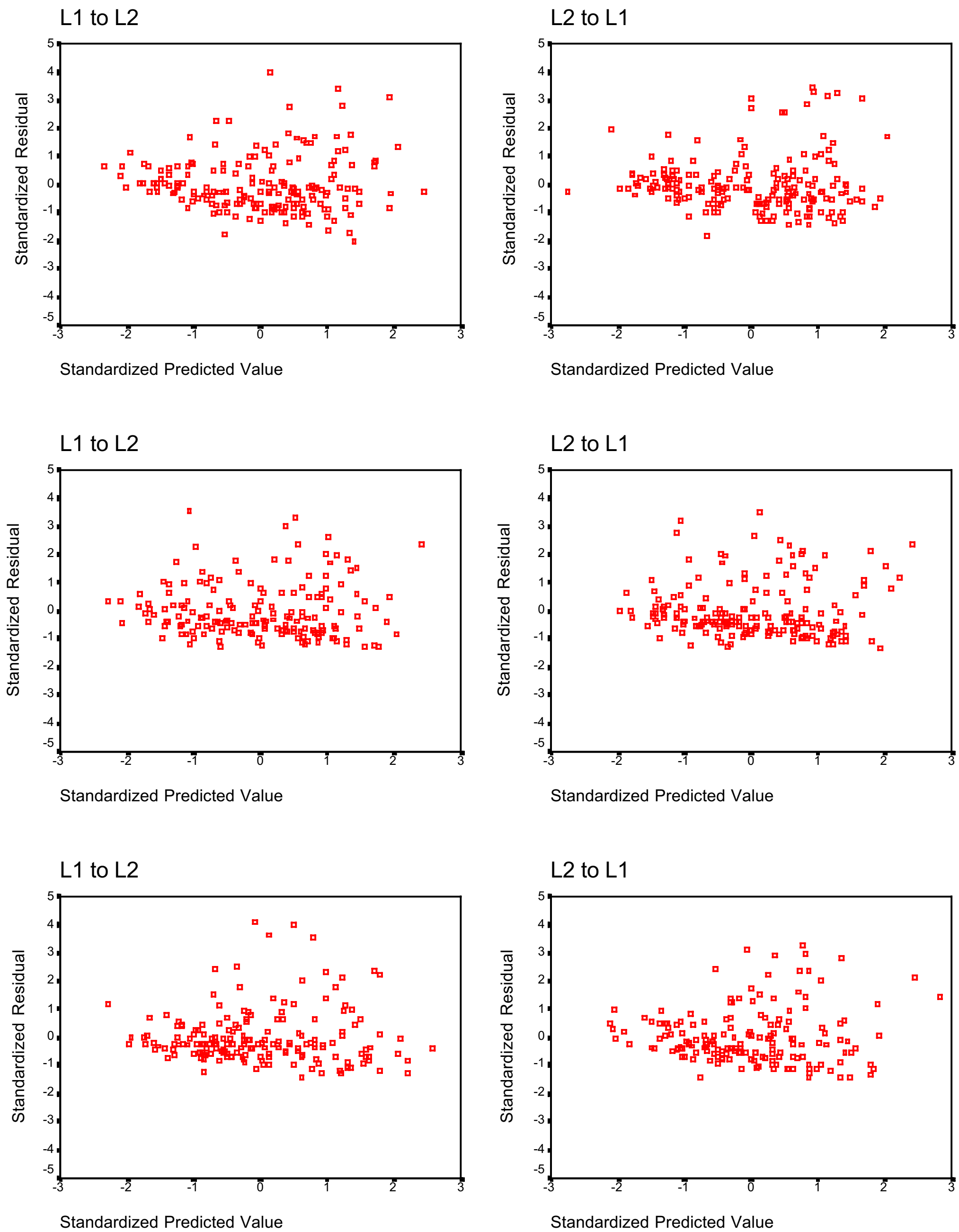


Figure 13. Scatterplots of predicted values against residuals for the two directions of translation from the conceptual salience analysis (taken from participants 1, 2, and 3 in the top, middle, and bottom panels, respectively).

Appendix D, continued

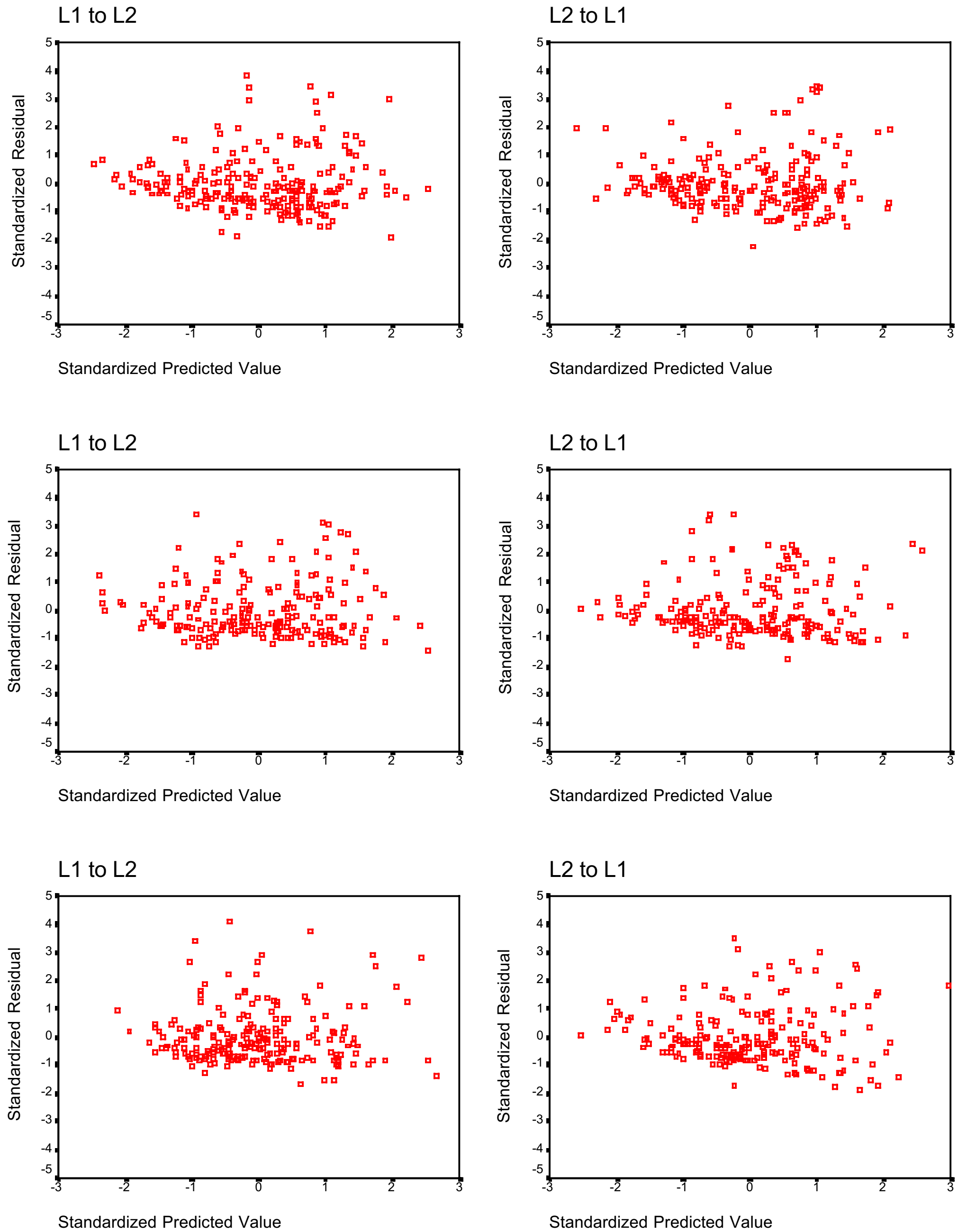


Figure 14. Scatterplots of predicted values against residuals for the two directions of translation from the ambiguity analysis (taken from participants 1, 2, and 3 in the top, middle, and bottom panels, respectively).

Appendix D, continued

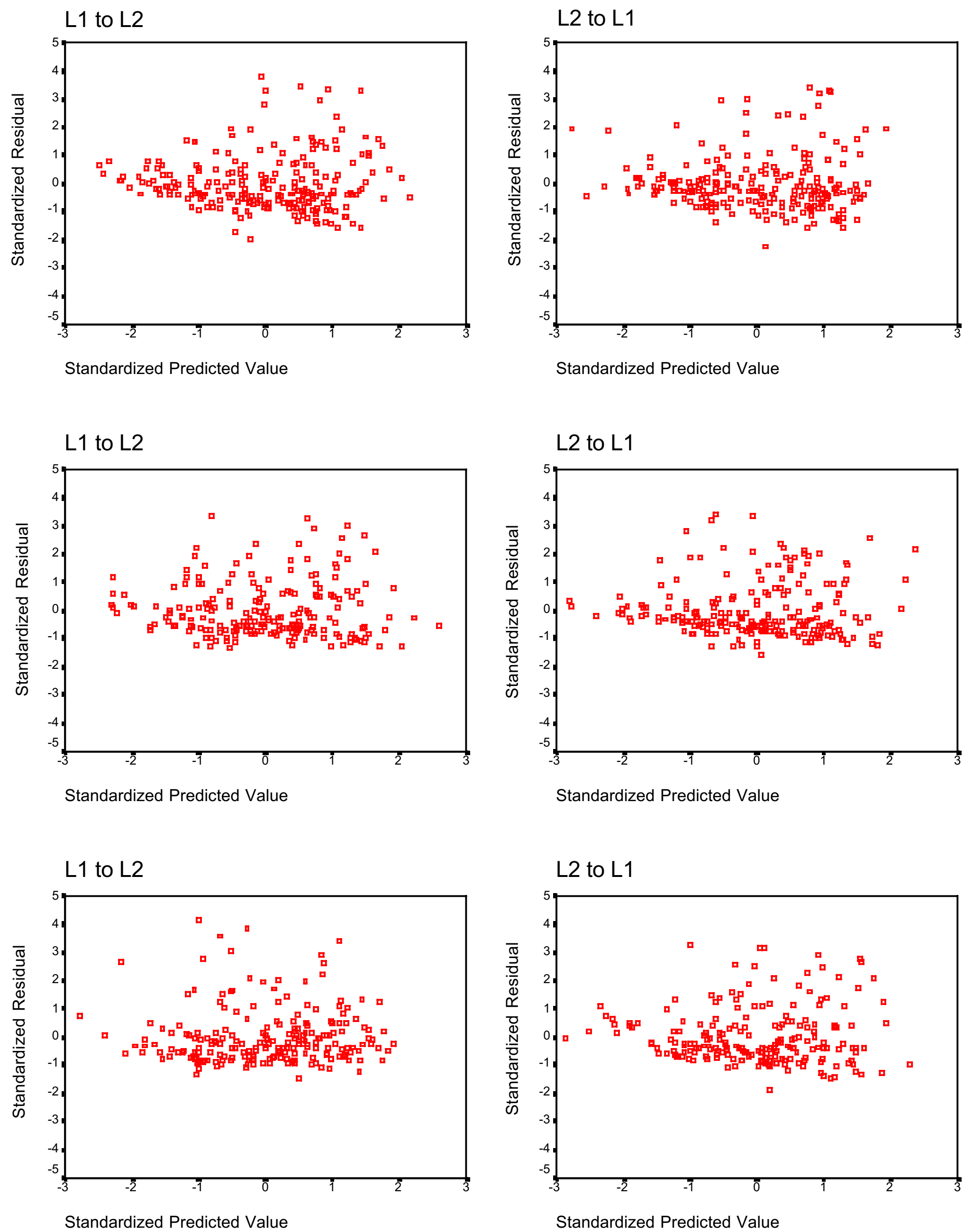


Figure 15. Scatterplots of predicted values against residuals for the two directions of translation from the cognate analysis (taken from participants 1, 2, and 3 in the top, middle, and bottom panels, respectively).

Appendix E

Instructions for Concreteness Ratings (Adapted from Spreen & Schulz, 1966)

Nouns may refer to persons, places, and things that can be seen, heard, felt, smelled or tasted or to more abstract concepts that cannot be experienced by our senses. The purpose of this experiment is to rate a list of words with respect to “concreteness” in terms of sense-experience. Any word that refers to objects, materials or persons should receive a high concreteness rating; any word that refers to an abstract concept that cannot be experienced by the senses should receive a low concreteness rating. Think of the words “cat” and “independence.” “cat” can be experienced by our senses and therefore should be rated as high concrete; “independence” cannot be experienced by the senses as such and therefore should be rated as low concrete (or abstract). Please rate the following words on a scale from 1 to 7, where 1 indicates least concrete and 7 indicates most concrete. If you have any questions, please ask me now. *Please use the full range of the scale in making your judgments.*

Appendix F

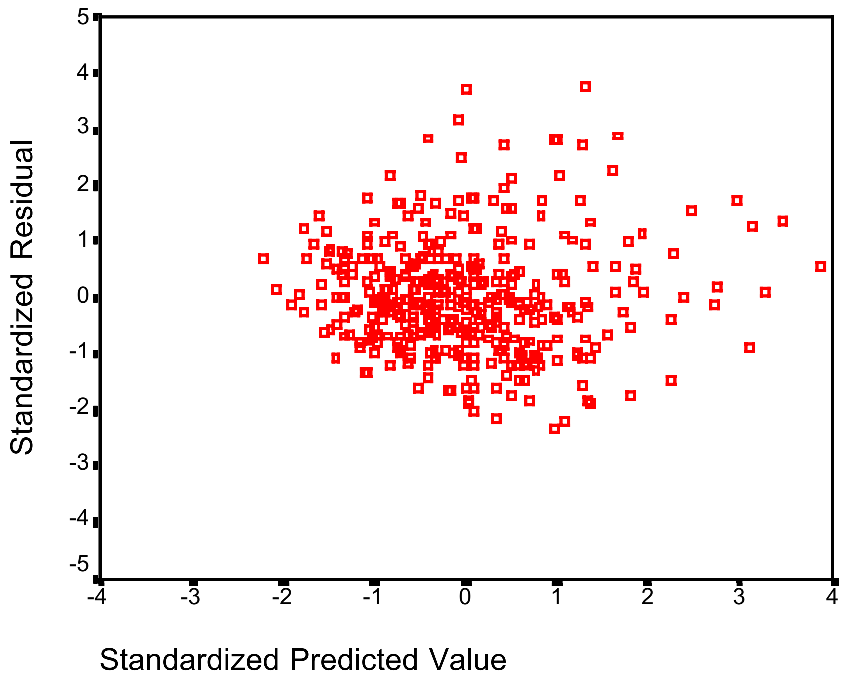
Instructions for Context Availability Ratings

(Adapted from Schwanenflugel, Harnishfeger, & Stowe, 1988)

Words differ on how easy it is to come up with a particular context or circumstance in which they might appear. It is easy to think of a context for the word "baseball" and "emotion," but it is much harder to think of a context for the word "inversion" and "sloop." Rate the following words on the ease with which you can think of a context for each word on a scale of 1 to 7, where 1 means "very hard to think of a context" and 7 means "very easy to think of a context." In the above examples, it is easy to think of a context for the word "baseball" and "emotion" (I immediately think of the context of the World Series for "baseball" and perhaps the context of falling in love for "emotion.") so they might receive a rating of 6 or 7. In contrast, it is rather difficult to think of a context for the words "inversion" and "sloop." (It is more difficult, but finally I think of atmosphere changes in Los Angeles for "inversion" and sailing in the Atlantic for "sloop.") so they might receive a rating of 1 or 2. Please use the full range of the scale in doing this task. If you are not sure how to rate a particular word, please make your best guess. If you have any questions, please ask me now.

Appendix G

Scatterplot of the Predicted Values against Residuals from Experiment 2



Vita

Natasha Tokowicz

Education

1997-2000	Ph.D., Psychology, The Pennsylvania State University (Major in Cognitive Psychology, Minor in Psychobiology)
1995-1997	M.S., Psychology, The Pennsylvania State University
1991-1995	B.A., Psychology, cum laude, University of Massachusetts, Amherst (Spanish Minor)

Publications

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (in preparation). *The development of lexical fluency in a second language.* Invited chapter to appear in the 2002 Issue of Second Language Research.

Kroll, J. F., & Tokowicz, N. (in press). The development of conceptual representation for words in a second language. In J. L. Nicol & T. Langendoen (Eds.), *One mind, two languages: Bilingual language processing.* Cambridge, MA: Blackwell Publishers.

Tokowicz, N., & Kroll, J. F. (2000). *Accessing meaning for words in two languages: The effects of concreteness and multiple translations in bilingual production.* Manuscript submitted for publication.

Tokowicz, N., & Kroll, J. F. (in preparation). *Priming interlanguage connections: Evidence for two routes to translation.* Manuscript in preparation.

Presentations

Tokowicz, N. (2000, March). *How do "glass" and "culture" differ?: Implications for bilinguals.* Poster presented at the Fifteenth Annual Graduate Research Exhibition, The Pennsylvania State University, University Park, PA.

Tokowicz, N., & Kroll, J. F. (1999, November). *Conceptual Access in Bilingual Translation.* Poster presented at the Fortieth Annual Meeting of the Psychonomic Society, Los Angeles.

Tokowicz, N., & Kroll, J. F. (1998, May). *Assessing the Role of Meaning in Bilingual Translation.* Paper presented at the Tenth Annual Meeting of the American Psychological Association, Washington, DC.

Tokowicz, N., & Kroll, J. F. (1998, April). *When is meaning accessed for second language words? It depends on what you mean!* Paper presented at the Forty-Third Annual Meeting of the International Linguistic Association, New York, NY.

Kroll, J. F., Michael, E., Elsinger, C., Tokowicz, N., & Miller, N. (1997, April). *Early stages of second language learning: The role of individual differences in acquiring L2 vocabulary.* Paper presented at the International Symposium on Bilingualism, Newcastle-upon-Tyne, United Kingdom.

Johnson, S. H., Hawley, E., Tokowicz, N., & Rosenbaum, D. A. (1996, November). *Virtual reaching to visual objects.* Poster presented at the Thirty-Seventh Annual Meeting of the Psychonomic Society, Chicago, IL.

Tokowicz, N., & Kroll, J. F. (1996, March). *Priming two routes to translation.* Paper presented at the Sixty-Seventh Annual Meeting of the Eastern Psychological Association, Philadelphia, PA.

Tokowicz, N. (1996, March). *Language Representation in the Bilingual: Evidence for Two Routes to Translation.* Poster presented at the Eleventh Annual Graduate Research Exhibition, The Pennsylvania State University, University Park, PA.

Kroll, J. F., Elsinger, C., & Tokowicz, N. (1994, November) *Priming interlanguage connections: Evidence for two routes to translation.* Paper presented at the Thirty-Fifth Annual Meeting of the Psychonomic Society, St. Louis, MO.

Selected Grants and Awards

2000	Department of Psychology Dissertation Support Award
2000	2nd Place Award, Social & Behavioral Sciences, 15th Graduate Research Exhibition
1999	College of the Liberal Arts RGSO Dissertation Support Grant
1998	Sigma Xi Grant-in-Aid of Research
1997	Sigma Xi Grant-in-Aid of Research
1997	Sigma Xi Pennsylvania State University Chapter Matching Research Grant Award
1997	Eileen Wirtshafter & Herschel W. Leibowitz Fund Travel Award