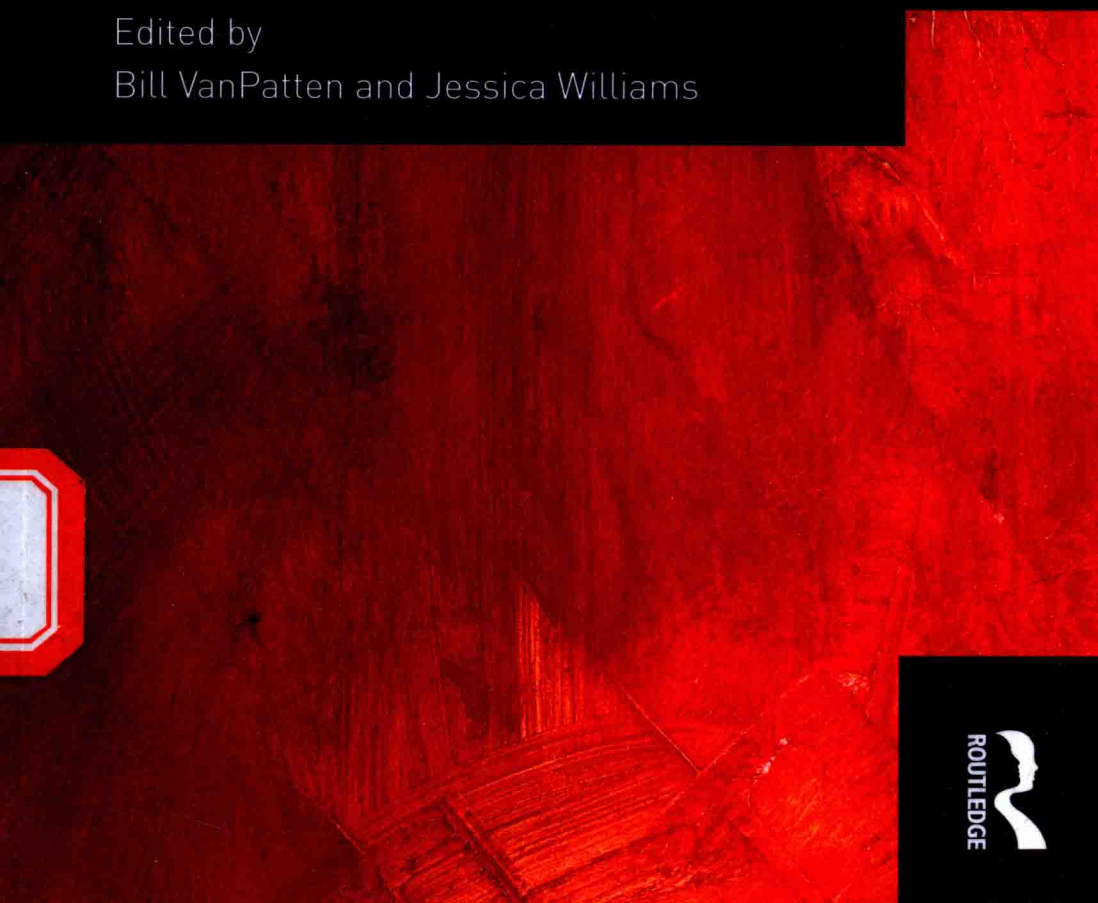
Second Edition

Second Language Acquisition Research Series

THEORIES IN SECOND LANGUAGE ACQUISITION

AN INTRODUCTION

Edited by
Bill VanPatten and Jessica Williams



ROUTLEDGE

THEORIES IN SECOND LANGUAGE ACQUISITION

An Introduction

Second Edition

Edited by Bill VanPatten and Jessica Williams

MICHIGAN STATE UNIVERSITY AND
UNIVERSITY OF ILLINOIS AT CHICAGO

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

This edition first published 2015
by Routledge
711 Third Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

[First edition published by Lawrence Erlbaum Associates, Inc. 2007]

Library of Congress Cataloging-in-Publication Data
Theories in second language acquisition : an introduction / Bill VanPatten and Jessica Williams, Michigan State University and University of Illinois at Chicago. — Second Edition.
pages cm. — (Second Language Acquisition Research Series)
Includes bibliographical references and index.
1. Second language acquisition. I. VanPatten, Bill, editor. II. Williams, Jessica, 1957– editor.
P118.2.T45 2015
418.0071—dc23
2014018009

ISBN: 978-0-415-82420-0 (hbk)
ISBN: 978-0-415-82421-7 (pbk)
ISBN: 978-0-203-62894-2 (ebk)

Typeset in Bembo
by Apex CoVantage, LLC

THEORIES IN SECOND LANGUAGE ACQUISITION

The second edition of the best-selling *Theories in Second Language Acquisition* builds on the strengths of the first edition by surveying the major theories currently used in second language acquisition (SLA) research, serving as an ideal introductory text for undergraduate and graduate students in SLA and language teaching. Each chapter focuses on a single theory, written by a leading scholar in the field in an easy-to-follow style—a basic foundational description of the theory, relevant data or research models used with this theory, common misunderstandings, and a sample study from the field to show the theory in practice. This text is designed to provide a consistent and coherent presentation for those new to the field who seek basic understanding of theories that underlie contemporary SLA research but will also be useful to researchers as a "quick guide" to theoretical work outside their respective domains.

Bill VanPatten is professor of Spanish and second language studies at Michigan State University.

Jessica Williams is professor of linguisics at the University of Illinois at Chicago.

Gass/Sorace/Selinker Second Language Learning Data Analysis, Second Edition (1998)

Mackey/Gass Second Language Research: Methodology and Design (2005)

Gass/Selinker Second Language Acquisition: An Introductory Course, Third Edition (2008)

PREFACE

This book focuses on a number of contemporary mainstream theories in second language acquisition (SLA) research that have generated attention among scholars. For several decades, the field of SLA has struggled with the nature of theories, what they are, and what would be an "acceptable" theory of SLA. Indeed, the present volume draws on one particular publication by Michael Long in a special issue of the *TESOL Quarterly* from 1990 devoted to the construction of a theory in SLA. In that article, Long discussed the nature of what a theory needs to be in SLA and also summarized the research to establish "the least" a theory of SLA needs to explain. We borrow from Long's article in our first chapter to outline the challenges to contemporary theories and list 10 observations that need to be accounted for on theoretical grounds.

One might ask why there are so many "competing" theories in SLA at this point. Why isn't there just one theory that accounts for SLA? What is it about SLA that invites a diffusion of theoretical perspectives? To understand this, one might consider the parable about the four blind men and the elephant. These sightless men chance upon a pachyderm for the first time and one, holding its tail, says, "Ah! The elephant is very much like a rope." The second one has wrapped his arms around a giant leg and says, "Ah! The elephant is like a tree." The third has been feeling alongside the elephant's massive body and says, "Ah! The elephant is very much like a wall." The fourth, having seized the trunk, cries out, "Ah! The elephant is very much like a snake." For us, SLA is a big elephant that researchers can easily look at from different perspectives. SLA is, after all, an incredibly complex set of processes, and if you have been introduced to the field via any of the excellent overviews of SLA, this most likely is your conclusion. Thus, researchers have grabbed onto different parts of the elephant as a means of coming to grips with the complex phenomenon. This does not mean, however, that researchers and scholars have

gone poking around SLA blindly and without thought; the present chapters should convince you otherwise. Unlike the blind men of our fable, researchers grasp that to understand the whole of SLA, they may need to concentrate on the smaller parts first. In the end, we may even need multiple complementary theories to account for different observed phenomena of SLA. As you complete the readings in your book, you might ask yourself, "Just what part of the elephant is each theory examining?"

The present book came about as a perceived need to have a comprehensive yet readily accessible set of readings for the beginning student of SLA. Each of us has taught introductory courses on SLA to students in TESOL and applied linguistics, and we have felt that a good introduction to theories is beneficial. At the same time, we know that it is easy for authors who don't work in a particular theory to reduce the theory to the point of students misinterpreting it or to misinterpret the theory themselves and pass on this misinterpretation to students. To this end, we decided that a collection of chapters written by the experts who work in the theories would best suit our needs as well as those of our students. We are pleased to present this volume for the beginning student of SLA.

Since the publication of the first edition of this book, the field has continued to develop, incorporating insights from theories and research methods from other fields. In response to some of these developments, we have added two new theories to the original set in the first volume. However, it is important to be clear that this book does not cover *all theories* of SLA. Notably, it does not cover theories that take "a social turn." The focus of the original book was on linguistic, psycholinguistic, and cognitive perspectives in SLA, and the second volume has maintained this focus. Since the publication of the first edition, there have been several fine books exploring alternative and, in particular, more social perspectives on SLA. We believe that they complement the current volume.

ACKNOWLEDGMENTS

Since its inception, this volume has been developed with the novice reader in mind—the beginning student of SLA who may not have much background in linguistics or SLA. Keeping that novice reader in mind has been a challenge for us and no less for the various contributors whose theories you will read here. The process of getting this volume into final form was long and demanded considerable effort on the part of the contributors to present some very complex notions in an accessible and consistent format. We know this often tried the patience of our authors. We took them away from their research and teaching duties to answer our numerous queries and revise their chapters, not once but, for most of the authors, now twice for the second edition. That they stuck with us to the end is a demonstration of their commitment and dedication to the profession and to its newest members. They have our heartfelt thanks. We also thank Megan Smith, who worked to format and finalize the manuscript before it went to the publisher. Finally, we thank the folks at Routledge for bringing this volume into the hands of the reader.

CONTENTS

1

INTRODUCTION

The Nature of Theories

Bill VanPatten and Jessica Williams

Almost everyone has heard of Einstein's Theory of Relativity. People have also heard of things such as the "Theory of Evolution" and "Atomic Theory." What is common to all these theories is that they are theories about what scientists call *natural phenomena*: things that we observe everyday. Theories are a fundamental staple in science, and all advances in science are, in some way or another, advances in theory development. If you asked scientists, they would tell you that the sciences could not proceed without theories. And if you ask applied scientists (such as those who develop medicines or attempt to solve the problem of how to travel from Earth to Mars), they would tell you that a good deal of their work is derived from theoretical insights.

Theories are also used in the social and behavioral sciences, such as psychology, sociology, and economics. As in the natural sciences, social sciences attempt to explain observed phenomena, such as why people remember some things better than others under certain conditions or why the stock market behaves the way it does.

In the field of second language acquisition (hereinafter SLA) research, theories have also come to occupy a central position. Some researchers, though by no means all, would even say that the only way SLA can advance as a research field is if it is *theory driven*. The purpose of the present book is to introduce the reader to certain current theories in SLA and provide a background for continued in-depth reading of the same. As a starting point, we will need to examine the nature of theories in general.

What Is a Theory?

At its most fundamental level, a **theory** is a set of statements about natural phenomena that explains why these phenomena occur the way they do. In the sciences, theories are used in what Kuhn (1996) calls the job of "puzzle solving."

By this Kuhn means that scientists look at observable phenomena as puzzles or questions to be solved. Why does the earth revolve around the sun and not fly off into space? Why are humans bipedal but gorillas knuckle-walkers? These are all questions about things that confront us every day, and it is the job of scientists to account for them.

In short, then, the first duty of a theory is to *account for* or *explain* observed phenomena. But a theory ought to do more than that. A theory also ought to make predictions about what would occur under specific conditions. Let's look at three examples: one familiar, the other two perhaps less so. In the early part of the 19th century, scientists were already aware of the presence of microorganisms in the air and water, and they had an idea about the connection between the organisms and disease. However, they had no idea of how they came into existence; indeed, belief in the spontaneous generation of these organisms was widespread. Disease was thought to be caused by "bad air." Careful experimentation by Louis Pasteur and other scientists demonstrated that microbes, though carried by air, are not created by air. Living organisms come from other living organisms. These discoveries led to the development of the *germ theory of disease,* which proposed that disease was caused by microorganisms. The acceptance of this theory had obvious important applications in public health, such as the development of vaccines, hygienic practices in surgery, and the pasteurization of milk. It not only could *explain* the presence and spread of disease, it could also *predict,* for example, that doctors who delivered babies without washing their hands after performing autopsies on patients who had died from childbirth fever would transmit the disease to new patients. Even more important, the same theory could be used to *connect* phenomena that, on the surface, appeared unrelated, such as the transmittal of disease, fermentation processes in wine and beer production, and a decline in silkworm production.

Now let's take an example from psychology. It is an observed phenomenon that some people read and comprehend written text faster and better than others. As researchers began to explore this question, a theory of individual differences in working memory evolved. That theory says that people vary in their ability to hold information in what is called working memory (defined, roughly, as that mental processing space in which a person performs computations on information at lightening speed). More specifically, the theory says that people vary in their working memory *capacity*: Some have greater capacity for processing incoming information compared with others, but for everyone, capacity is limited in some way. Initially used to account for individual differences in reading comprehension ability in a person's first language, the theory also accounts for a wide range of seemingly unrelated phenomena, such as why people remember certain sequences of numbers and not others, why they recall certain words that have been heard, why people vary on what parts of sentences they remember best, why certain stimuli are ignored and others attended to, and why some students are good note takers and others are not. A theory of working

memory, then, allows psychologists to unify a variety of behaviors and outcomes that on the surface level do not necessarily appear to be related. There are even attempts to apply the theory of SLA to explain why some people learn faster and better than others.

Let's take a final example, this time from language. In one theory of syntax (sentence structure), a grammar can allow movement of elements in the sentence. This is how we get two sentences that essentially mean the same thing, as in the following:

(1) Mary said what?
(2) What did Mary say?

In this particular theory, the *what* is said to have moved from its position as an object of the verb *said* to occupy a place in a different part of the sentence. At the same time, this theory also says that when something moves, it leaves a hidden *trace*. Thus, the syntactician would write (2) like (3):

(3) What$_i$ did Mary say t_i?

In (3) the *t* stands for the empty spot that the *what* left and the *i* simply shows that the *what* and the *t* are "co-indexed"; that is, if there happens to be more than one thing that moves, you can tell which trace it left behind.

To add to the picture, the theory also says that *t*s, although hidden, are psychologically real and occupy the spot left behind. Thus, nothing can move into that spot and no contractions can occur across it. Armed with this, the syntactician can make a variety of predictions about grammatical and ungrammatical sentences in English. We might predict, for example, that (4) is a good sentence but (5) is bad and not allowed by English grammar:

(4) Should I have done it?
(5) Should I've done it?

The reason for this is that *should* has moved from its original spot and left a *t* behind, as illustrated in (6):

(6) I should have done it. → Should$_i$ I t_i have done it?

At the same time, the syntactician would predict restrictions on the contraction of *want to* to *wanna*. Thus, (7) is fine because there is no trace intervening where a contraction wants to happen:

(7) Who$_i$ do you want to invite t_i to dinner? → Who do you wanna invite to dinner?

All English speakers would agree, however, that (8) is awful:

(8) *Who do you wanna take Susie to the prom next month?

You could probably work this out yourself, but the reason (8) sounds bad is that the *who* has moved and has left behind a *t* that blocks a possible contraction. Compare (7) and (8) redone here as (9) and (10):

(9) Who$_i$ do you want to invite t_i to dinner? → Who do you wanna invite to dinner?
(10) Who$_i$ do you want t_i to take Susie to the prom next month? → *Who do you wanna take Susie to the prom next month?

Be careful not to pronounce *wanna* like *want tuh*; *want tuh* is not a contraction and is merely the schwawing of the vowel sound in *to*. *Want tuh* sounds OK in sentence (8) precisely because it is not a contraction.

Thus, the theory unifies constraints on contractions with modals (*should, would, will, may, might*), with auxiliaries (*do, have*), with copular verbs (*be*), with the verb *want,* and with pronouns (*I, you, he,* and so on). It makes predictions about good and bad sentences that perhaps we have never seen or heard, some of which—like silkworms and beer—don't seem to have much in common.

To summarize so far, a theory ought to account for and explain observed phenomena and also make predictions about what is possible and what is not. In addition, most theories—good ones, that is—when accounting for and predicting things, also tend to unify a series of generalizations about the world or unify a series of observations about the world. In the brief view we had of syntactic theory, the few generalizations made about how syntax works unify a variety of observations about contractions and not just contractions with *should*. All contractions conform to the generalizations.

For SLA, then, we will want a theory that acts like a theory should. We will want it to account for observable phenomena (something to which we turn our attention later in this chapter). We want it to make predictions. And, ideally, we want it to unify the generalizations we make as part of the theory. In other words, we want a single theory to bring all of the observed phenomena under one umbrella. Whether this is possible at this time has yet to be determined and is something that this book will explore.

What Is a Model?

Many people confuse theories and models. A **model** describes processes or sets of processes of a phenomenon. A model may also show how different components of a phenomenon interact. The important word here is *how*. A model does not need to explain *why*. Whereas a theory can make predictions based on generalizations,

this is not required of a model. The problem is that in the real world—and in SLA as a research discipline—this distinction is not always maintained. You will find as you read further in the field that researchers often use *model* and *theory* interchangeably. Thus, although in principle it would be a good idea to distinguish between these two terms as they do in the natural sciences, in practice many of us in SLA do not do so.

What Is a Hypothesis?

Distinct from a theory, a **hypothesis** does not unify various phenomena; it is usually an idea about a single phenomenon. Some people use theory and hypothesis interchangeably, but in fact, they are distinct and should be kept separate. In science, we would say that a theory can generate hypotheses that can then be tested by experimentation or observation. In psychology, for example, there are theories regarding memory. You may recall the theory about working memory and capacity discussed earlier. The theory says (among many other things) that working memory is limited in capacity. This means that people can pay attention to only so much information at a given time before working memory is overloaded. The theory also says that there are individual differences in working memory and how people use what they have. Some people have *X* amount of working memory capacity as they attend to incoming information, whereas others have more or less. A hypothesis that falls out of this, then, is that working memory differences among individuals should affect reading comprehension: Those with greater working memory capacity should be faster readers or should comprehend more. This is a testable hypothesis. We ought to add here that the only valuable hypotheses for a theory are those that are testable, meaning some kind of experiment can be run or some kinds of data can be examined to see if the hypothesis holds up. Another example of a hypothesis comes from SLA: the Critical Period Hypothesis. This is a theory in neurolinguistics that states that at an early age, the brain begins to specialize; specific brain functions become increasingly associated with specific areas of the brain. In addition, some functions may be developmentally controlled; that is, they turn on and, more important for language learning, turn off at specific points in development. The Critical Period Hypothesis is a direct consequence of this theory. It states that the ability to attain native-like proficiency in a language is related to the initial age of exposure. If language learning begins after a certain age (and there is a considerable controversy over what this age is as well as whether there even is a critical period—see the various papers in Birdsong, 1999), the learners will never reach a level of proficiency or competence comparable to a native speaker's. A corollary to this hypothesis is that language learning ability declines with age after this point. Again, both of these are testable hypotheses. Recall that earlier we said we wanted a theory to make predictions. Predictions are actually hypotheses. When we make a prediction based on a theory, we are in effect making a hypothesis.

These definitions about theories, models, and hypotheses are important because in everyday speech, we may use the term *theory* in a way not intended in science. For example, one might hear in a disparaging tone that something is "just a theory." In science, the phrase "just a theory" makes no sense, as all work is theoretically driven. What is more, the term *theory* has often been politicized to denigrate particular theories (e.g., evolution) so that "just a theory" becomes a way of dismissing something that has scientific rigor but runs against some other set of beliefs. Finally, in movies and other nonscientific situations, one often hears the term *theory* used to mean "an idea" or a "hypothesis." A detective trying to solve a crime might say, "I have a theory about the killer," when that detective means, "I have an idea about the killer." We cannot, of course, rid everyday speech of how it uses certain words. Our point in bringing up the everyday use of *theory* is to make sure that the reader understands the term as it is used in this book.

Constructs

All theories have what are called **constructs**. Constructs are key features or mechanisms on which the theory relies; they must be definable in the theory. In the theory about disease transmission, *germ* is a construct. In the theory about working memory, *capacity* is a construct; and in the theory about syntax, a *trace* is a construct.

In evaluating any theory, it is important to understand the constructs on which the theory relies; otherwise, it is easy to judge a theory one way or another—that is, as a good or bad theory—without a full understanding of the underpinnings of the theory. For example, without an understanding of the construct *germ*, it would have been easy to dismiss germ theory. But given that the construct *germ* was easily definable and identifiable, dismissal of germ transmission and diseases was not so facile. To fully understand something like Relativity, one must have a thorough grasp of the constructs *time, space,* and others.

In SLA, we find an abundance of constructs that are in need of definitions. For example, take the term *second language acquisition* itself. Each word is actually a construct, and you can ask yourself, "What does *second* mean?" "What does *language* mean?" and "How do we define *acquisition*?" In SLA theorizing, most people use the term *second* to mean any language other than one's first language. It makes no difference what the language is, where it is learned, or how it is learned. This suggests, then, that any theorizing about SLA ought to apply equally to the person learning Egyptian Arabic in Cairo without the benefit of instruction as to the person learning French in a foreign language classroom in the United States. By defining *second* in an all-encompassing way, it has an effect on the scope of the theory. If the construct *second* were not defined this way, then it would have limited scope over the contexts of language learning. For example, some people define *second* language to refer to a language learned where it is spoken (e.g., immigrants learning English in this country, an American learning Japanese in Osaka), whereas *foreign* is used to refer to situations in which the language is not

spoken outside of the classroom (e.g., German in San Diego, California). Thus, if *second* were defined in the more restricted way, a theory of SLA would be limited to the first context of learning.

The term *language* is deceptively simple as a construct, but have you ever tried to define it? Does it mean speech? Or does it mean the rules that govern speech production? Or does it mean the unconscious knowledge system that contains all the information about language (e.g., the sound system, the mental dictionary, syntactic constraints, rules on word formation, rules on use of language in context)? Thus, any theory about SLA needs to be clear on what it means by *language*. Otherwise, the reader may not fully grasp what the theory claims, or worse, misinterpret it.

In summary, here are key issues discussed so far:

- Theories ought to explain observable phenomena.
- Theories ought to unify explanations of various phenomena where possible.
- Theories are used to generate hypotheses that can be tested empirically.
- Theories may be explanations of a *thing* (such as language) or explanations of *how* something comes to be (such as the acquisition of language).
- Theories have constructs, which in turn are defined in the theory.

Why Are Theories and Models Either Good or Necessary for SLA Research?

We have explored what theories are but only obliquely addressed why they might be useful. Certainly, they help us to understand the phenomena that we observe. Consider again the Critical Period Hypothesis. It has often been observed that speakers who begin the process of SLA later in life usually have an accent. A theory about the loss of brain plasticity during natural maturation may help explain this phenomenon. The same theory might predict that learners who begin foreign language study in high school will be less likely to approach a native-like standard of pronunciation than those learners who have access to significant amounts of target-language input much earlier in life. These kinds of predictions have clear practical applications; for example, they suggest that foreign language learning should begin at a young age.

Let's look at another concrete example. In one theory of SLA, producing language (usually called *output*) is considered an important element in structuring linguistic knowledge and anchoring it in memory. In another theory, in contrast, output is considered unimportant in developing second language knowledge. Its role is limited to building control over knowledge that has already been acquired. These differences in theory would have clear and important consequences for second language instruction. In the first case, output practice would have a significant role in all aspects of instruction. In the second case, it would be most prominent in fluency practice.