INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

University of Alberta

Second Language Acquisition and Comprehension Mechanisms:

The Problem of Russian Inflectional Morphology

by

Gregory Thomson 

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree Doctor of Philosophy

in

Psycholinguistics

Department of Linguistics

Edmonton, Alberta

Fall 2000

University of Alberta

Library Release Form

Name of Author: Gregory Edward Thomson

Title of Thesis: Second Language Acquisition and Comprehension Mechanisms: The Problem of Russian Inflectional Morphology

Degree: Doctor of Philosophy

Year this Degree Granted: 2000

Gregory E Thomson

Box 42038 Millbourne RPO
Edmonton, Alberta, T6K 4C4

Date submitted: May 30, 2000

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Second Language Acquisition and Comprehension Mechanisms: The Problem of Russian Inflectional Morphology** submitted by Gregory Thomson in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in *Psycholinguistics*.

.....
Dr. G. Libben

.....
Dr. B.L. Derwing

.....
Dr. G.D. Prideaux

.....
Dr. L.M. Stanford

.....
Dr. L. Ranta

.....
Dr. B. Van Patten
(Spanish, Italian and Portuguese,
University of Illinois at U-C)

Date: May 25, 2000

Abstract

Acquiring grammatical form is typically considered central to second language acquisition (SLA), and assumed to involve the development of an internal grammar, i.e., a mental apparatus that characterizes the form of sentences in the second language (L2). An alternative assumed here is to view linguistic forms as processing triggers and SLA as learning to detect and react to such triggers (and then to produce comparable ones).

Inflectional cue systems are inherently complex, combining lexical cues and grammatical cues in phonological words. They are complex in their formal nature, in the sets of processes that they trigger, and in the relationships between the two. Thus, it should not be surprising that L2 inflection is often not readily acquired.

This dissertation first sketches a broad picture of language processing, next explores the nature of inflectional morphology, and then proceeds to an experimental investigation into whether various categories of inflectional morphology are active or inert in the comprehension systems of ninety-one adult learners of Russian as a second language, representing two developmental levels (along with twenty-five native Russian control participants). Sensitivity to inflectional errors (and by implication, to particular features of inflectional form) was investigated under three conditions: 1) meaning-oriented listening; 2) form-oriented listening; 3) free examination of printed sentence.

Evidence emerges that systematic acquisition of inflectional morphology does take place in learners of L2 Russian. However, the developmental process would appear to be so protracted that complete acquisition may be a rarity. Explicit metalinguistic knowledge of L2 grammar may allow some learners to simulate many features of nativelike production when the relevant strictly linguistic developments may still be years away. Implications are discussed in relation to various SLA models and research approaches.

For Angela

Acknowledgements

Many have helped me along the path to this dissertation. There was Plotkins, my fifth grade teacher in Van Nuys, who encouraged me by telling my mother of how she was impressed with my ability to understand grammar concepts. And there was my mother who passed the word on to me (along with a fair bit more). And Mr. Hatch, my sixth grade teacher who encouraged me to pursue science. Mr. Injijikian, my eleventh grade French teacher, who gave me the idea that what a language learner *does* affects what s/he *learns*. Then there were my teenage Carrier Indian friends who at various locations among the jackpines of the upper Dean River plateau showed me that French and English are more or less dialects of a single language. We had fun. An amateur linguist/missionary, Vaughn Collins, confirmed my suspicion that languages were more amazing than I had realized, and planted in me seeds of interest in linguistics.

Next began many wonderful friendships with members of the Siksika (Blackfoot) Nation. Those who especially encouraged me in my efforts to become part of their speech community included Bernard Tailfeathers, Bernard Bigsnake, and Jack Bigeye (who took me as a son). During that period, various missionaries passed on to me their sense of wonder at the beauty of linguistic and cultural diversity. This included especially Alan Franz, Alban Douglas and Howard Klaasen.

The people mentioned so far, and others, created in me a desire to know more about language. Gradually that desire had grown into a highly volatile mass. Next came the spark, in the person of Don Frantz. In 1967, he was a Ph.D. student at the University of Alberta. I thank him for many highly charged afternoons, discussing language, linguistics, and Blackfoot, in the summer of 1967, and for a number of years following. He has certainly been one of the most influential people in my life.

Next I wish to thank my early teachers in the Summer Institute of Linguistics (SIL) at the University of North Dakota. I needed to be around people who were excited about language, and I found many such people there in the summer of 1969. I am especially grateful to Dick Pitmann, George Cowan, and Randy Speirs, along with others, such as Ed Moser, Becky Moser, Warren Harbeck, Hank Bradley, John Crawford, Peter Landerman, Jean Donaldson, Helen Neiswander, and one or two others whose names fail me. From the following two summers, I would add the names of Paul Melema, Ken Gregerson, and John Daly. Thanks to all who helped make me a linguistics fan in my early adulthood.

There are also many I could thank from among my early colleagues in the North America Branch of the Summer Institute of Linguistics, but sticking to those who especially encouraged me in linguistics, I must mention Constance Naish and Wayne Leman (two of the dearest friends of my life) as well as my first SIL director, Irvine Davis.

My early desire to help language learners succeed was also rooted in SIL relationships, again Wayne and Elena Leman, Constance Naish, and also Fred Miska, Vic Monus, Curtis Cook, Randy Spiers (again), Ann Speirs, Dean Saxton, Lee Ballard, and my wife, Angela. I thank Tom and Betty Sue Brewster for giving me some tools and enthusiasm for helping language learners and Don Larson for the inspiration behind much of that. I also owe thanks my earliest language learning consultees, Wolf Seiler and (again) Fred Miska (and also Donna Gardiner and Barbara Allen, who were an indirect encouragement at that point, and later a direct encouragement).

Then came the North America Branch SIL Language Learning Club of 1979-1980. I actually considered dedicating this dissertation to Roger, Ruth, Meggie, Pierre, Ben, Elma, Lana, Carol, Constance, Gillian, Dan and Jeannie. And thanks to others who contributed to the excitement of those months, including Gene Burnham, Vic Monus (again), Dean Saxton (again), Randy Speirs (again and again), Roger Gilstrap, Scott Palmer, Lynanne Palmer, Dick Bartsch, and probably others I'm forgetting. Later important consultees were Ross and Marianne Amy, and Dave and Kathy Rising, who led me to start rethinking the way I went about helping language learners.

I would also like to thank some very special language learning consultees in Pakistan: Dave, Susan, Andie, Ellie, Harold, Lillian, Niel, Val, Raylene, Rosemary, Ann, and especially, Mike and Joan. How much we learned together! Also thanks to Austin Hale who provided occasional reminders during those years that there is little that is more exciting than language. Also thanks to Warren and Jessie Glover for their intellectual and spiritual stimulus at that time. Special thanks to Victor Dean who first nurtured Angela and me into the Paksitani-Urdu speech community.

Thank you to those who then encouraged me to go to grad school (in early 1990). This includes Al Pence, Don Gregson, David Ross, Paul Farncombe, Les Bruce, Stephen Levinsohn, Bob Dooley, and John Oller.

Thanks to my special teachers at the University of New Mexico. Those who stand out most in my mind are John Oller, Joan Bybee, Larry Gorbet, Otto Santa Ana, Vera John-Steiner, and my favourite teacher of all time, Jean Newman.

Thanks to Carol Orwig, former SIL international language learning coordinator, for many deep, long, and helpful discussions of SLA issues throughout the 1990s.

That brings me relatively close in time to this dissertation. My supervisor, Gary Libben made me want to study with him from the first moments I heard him lecture at the University of Calgary in the fall of 1990. My co-supervisor, Bruce Derwing, and the chair of my examining committee, Gary Prideaux (both of whose acquaintance dates back to 1970) encouraged me to apply for admission to the Ph.D. program when I was feeling lost. They went on to enrich my life in the Linguistics Department. Thanks to John Hogan and Terry Nearey for wonderful conversations in the hall.

Thanks to Ross and Marianne Amy for their role in my life during the transition into the Ph.D. program. What memories!

Warren Harbeck's input was also providential at that same time.

Thanks to the Social Sciences and Humanities Research Council of Canada for three years' support (Award # 752-95-1523). Thanks to many others who have shared generously with Angela and me for these thirty-plus years.

Thanks to Marina Blekher, my fellow-Ph.D. student, the same age as my oldest son, and the one who first nurtured Angela, Ambrose, Chad and me into the Russian speech community. Thanks to Larissa Day who also sweetened our early Russian learning in Edmonton. Thanks to Svetlana from the Ukraine, whose last name we never knew, but who got us to use Russian and only Russian.

And thanks to those who followed up once we were in Russia, especially Ira Sherbina who played an enormously special role in nurturing us into Russia, along with Dima and Kira Ilyin, who enriched the latter part of our time there, and gently helped us to keep growing right up to our final hours there. Thanks to John and Debbie Clifton and Daniel and Elsbeth Angst who helped us into Russia in the first place. That reminds me that I should also give special thanks to the participants of my 1994 St. Petersburg language learning workshop, and my special consultees in southern Russia in 1997, in particular, Duncan, Mike, Perry, and Shelly. We were trying to figure out what makes Russian so hard for adult learners. Hopefully, we made some progress.

Thanks to my Russian language consultants, including my summer student helpers, Natalja Malijevskaja and Svetlana Voronina in Rostov-on-Don. Thanks also to Valentina Romanova at Rostov State University, who helped me in my first efforts to explain my research goals in Russian, and to create an experiment. Around the same time, Irina Tishchenko at Rostov Pedagogical University graciously invited me to audit her course on Russian language teaching methodology. That was my first experience in a Russian university and it was more fun than I could put into words (apart from a few extremely embarrassing moments resulting from my low level of Russian).

As my experimental work took its ultimate form, my primary consultant was Tamara Ivanova at St. Petersburg State University. Her assistance was both invaluable and indispensable. Many others in the St. Petersburg academic community were extremely helpful, supportive, and always gracious. I also received great help and encouragement from my teachers at the Swiss Center Russian Language School in St. Petersburg, Marina Abashina and Tatjana Filosofova. A great place for Russian classes, by the way.

Most recently, I have been feeling especially indebted to my supervisors, Gary Libben and Bruce Derwing, for helping me to improve the quality of many parts of this dissertation. They also were generous in allowing me liberty to stray a bit from the more usual format of experimental dissertations. The additional subject matter was important to me. It may also give a greater number of my friends and relatives some things to enjoy than would have been the case had I only written Chapter 3! Thanks also to the other committee members, Gary Prideaux, Lois Stanford, Leila Ranta, and my external reader, Bill VanPatten for helpful input.

Thanks to my parents for being my friends all these years, along with my brothers, Randy and Geoff. Thanks to my kids for being part of a fun and funny family with more books than anything else, as we traipsed around the world. (Little Carmen once summed up my interests as "babies, books and coffee".) Cam, Ron, Carmen, Holly, Ambrose and Chad are a special bunch of kids (forever). They are the very ones I would have asked for if I had been given the choice. Thanks to each of them for their parts in my life to this moment. Ambrose, Chad and their mother had no small role in the events directly leading to this dissertation. The biggest thanks go to Angela for her support and love through this thirty-some year process of getting my linguistic education, while we rode herd on all those kids together.

Finally, I thank the Creator of life for letting me know that reality means something. He also has let me be reborn into three different speech communities in addition to my native one. He created the splendour of human linguistic diversity. As with everything else, He did it for good reason. I thank Him for my life, with all its textures and seasons—for each part as much as for each other part. I thank Him for the lives of those I love and the lives those who love me, for the lives of those mentioned above, and the lives of many others who came to mind as I wrote, and more who didn't. I even thank Him for your life, whoever you might happen to be, reading this right now. "And God saw every thing that he had made, and, behold, it was very good."

Contents

List of Figures

List of Tables

List of Abbreviations

1pers	first person
2pers	second person
3pers	third person
acc	accusative
fem	feminine
impf	imperfective
instr	instrumental
loc	locative
masc	masculine
nom	nominative
perf	perfective
pl	plural
pres	present
sg	singular

0. Introduction: Fossilizable Phenomena in Comprehension

One of the central puzzles that Selinker discussed in his foundational "Interlanguage" paper (1972) was *fossilization*. A recent glossary of second language acquisition (SLA) terminology defines fossilization simply as "cessation of learning" (Gass & Selinker 1994, p. 332). The observable evidence of this cessation of learning is the presence of *fossilized forms*, that is, recurrent features of the nonnative speaker's spoken output which distinguish it from the spoken output of native speakers. Selinker discussed the persistence of these features, noting that they "remain as potential performance, reemerging in the productive performance of an IL [interlanguage] even when seemingly eradicated" (Selinker, 1972, p. 215). The areas in which this cessation of learning occurs, according to Selinker's (1972) discussion, include word order patterns and grammatical morphemes, the latter involving function words (such as English articles) and inflectional morphology (such as English plural marking).

Although the term fossilization will not commonly be encountered in the chapters that follow, this dissertation is in fact concerned with what Selinker (1972) referred to as fossilizable linguistic features. In everyday terms, it is about why some features of second languages (L2s) appear to be extremely difficult for adult learners to acquire, contributing to what Schachter (1990) termed the *incompleteness* of SLA. The linguistic domain I selected for exploration is inflectional morphosyntax, a classical area of incompleteness in SLA. Having made that choice, I next chose Russian as the L2 for this investigation, since it presents learners with a rich, multifaceted inflectional system. The data I chose to gather and examine are data related to the employment of inflectional form during listening comprehension, since I argue that the functions of inflection are first and foremost comprehension functions. If this is correct, then the first place to look for inflectional learning, or for the alleged arrest of inflectional learning, is in the realm of comprehension processes.

Fossilization was not originally studied as a feature of L2 comprehension. Rather, Selinker took the observable evidence of fossilization to be the presence of fossilized forms in a nonnative speaker's spoken output, stating that the empirical data that are relevant to understanding second language learning in general are attempts by adults "to express meanings, which [they] may already have, in a language which [they are] in the process of learning" (Selinker, 1972, p. 210). On this view, patterns of expression in spoken output are the visible manifestation of language learning. Therefore, the cessation of learning will also be detectable in these patterns, in particular, when the patterns stop evolving in the direction of nativelikeness. Fossilizable features may give way to apparent gains, only later to have the territory reclaimed by the earlier nonnative patterns. Thus Selinker noted that fossilizable features "remain as potential performance, reemerging in the productive performance of an IL [interlanguage] even when seemingly eradicated" (Selinker, 1972, p. 215). He suggested that "perhaps the most crucial fact... that any adequate theory of second-language learning will have to explain is the regular reappearance or reemergence" of these fossilized speech patterns (p. 216).

Selinker's emphasis on spoken production patterns as the visible manifestation of learning (including incomplete learning) can be seen in various proposals since that time. It is seen in the "box and arrow" models of Krashen (1982), Bialystok (1978), Ellis (1990), VanPatten (VanPatten & Cadierno, 1993), Gass (1988, 1997) Skehan (1998) and others.[1]

All of these models begin with input (some subset of the target language

[1] More recently, VanPatten (1996) has broken with this tradition, making the result of acquisition not output, but rather a developing language system.

utterances that the learner hears), proceed to internal processing (which, depending on the model, might or might not involve innate linguistic mechanisms, attention to form, conscious cognitive processes, etc.) and ultimately end in spoken output as the fruit of learning. The same emphasis is seen in work by McLaughlin (1990; 1987, Chapter 6), Schmidt (1990, 1992), Meisel, Clahsen and Pienemann (1981) and Pienemann (1999). Some of these researchers make reference to the development of speech production mechanisms, while others simply take spoken production as a reflection of underlying competence (which is neutral with regard to the modalities of comprehension and production—Chomsky, 1965). In all of these approaches, an internal change takes place in the learner, constituting learning and learning expresses itself preeminently in spoken production.

More recently there has been growing interest in the L2 development of comprehension mechanisms (see especially Carroll, 1999).[2] This includes work on L2 speech perception (see Flege, 1991; Leather & James, 1996), syntactic parsing (Juffs, 1998a,b; Juffs & Harrington, 1995, 1996) and work on the role of grammatical and semantic cues in determining agent and patient (see MacWhinney, 1997). Work related to the bilingual mental lexicon (see Kroll & de Groot, 1997; Libben, 2000) can also be included here, insofar as it deals with learning and with lexical access during comprehension.

The presence of nonnative features in the L2 spoken production of an individual should alert us to the existence of nonnative features in the comprehension system of the same individual. This is because it is hard to imagine the side-by-side existence in a single L2 user of a comprehension system that is nativelike, for example, in that it expects to find definite articles in context X and a production system that is nonnativelike in that it does not produce definite articles in the same context X. At least if such a situation were to exist, the L2 user ought to be highly aware of it, unless there were some mechanism to prevent his or her comprehension system from processing the output of his or her production system.

Once we are clear that what Selinker (1972) termed "fossilizable linguistic phenomena" must exert their effect on the comprehension system as well as in the production system, we need to ask, how does fossilization manifest itself in the L2 comprehension system? Of course, to answer that question, it is necessary to look at L2 comprehension processes. In fact, we should feel insecure about any picture of learning derived from focussing exclusively on output. These output-centered pictures may subtly include approaches to SLA framed in terms of which potential productions are "allowed by the grammar". Such pictures must at best be incomplete. And it might turn out that in the presence of a fuller picture of L2 development, including comprehension as well as production, the stubborn reemergence of the offending features would become less mysterious.

We cannot be sure that the development of the SLA field would not have taken a different course after 1972 if equal attention had been paid to comprehension and production at all levels, including the level of grammatical details. It may seem that production is easier to study than comprehension. From the standpoint of experimental research, one could argue that the opposite is the case. At least the volume of experimental

---

[2] There is a larger body of work dealing with factors that affect overall success in listening comprehension. See Rubin (1994) for a review. In general such work is not concerned with the more specific mechanisms of on-line L2 comprehension and their development, but rather looks at comprehension in more global terms, for example, measuring the ability of L2 learners to recall the content of sentences or texts when listening conditions are altered in various ways (e.g., changes in speech rate, changes in contextual support, etc.).

work on L1 speech comprehension in general is greater than the volume of work on speech production. Learning may be reflected in production data as well as comprehension data. However, it is possible that comprehension data are "cleaner" than production data in the following sense. Learners can apply metalinguistic knowledge to planning utterances and may feel motivated to do so in order to win approval from a teacher or in order to give the appearance of greater nativelikeness.

In the case of listening comprehension, there is far less scope for applying metalinguistic knowledge and less motivation to do so. This makes it likely that the pattern Selinker discusses of "eradication" followed by "reemergence" will not be as characteristic of comprehension as it is of production, since gains in comprehension will more likely be genuine linguistic gains.

As noted, the domain I have chosen for investigation is L2 Russian inflectional morphology, which appeared to promise a rich set of fossilizable inflectional phenomena. The eradication of errors from L2 spoken Russian is a commonly emphasized pedagogical goal (Rifkin, 1995). Out of the nine notorious error types considered by Rifkin (1995), seven could be called errors of inflectional morphosyntax. Among learners of L2 Russian in Russia, "the endings" (that is, inflectional forms) are the most commonly bemoaned source of frustration. Nonnative inflectional form in spoken output is observed even in the speech of relatively fluent nonnative speakers such as news correspondents. In any case, even apart from pedagogical concerns, Russian inflection provides a promising domain for examining nonnativelike patterns of L2 speech that are difficult to "eradicate", where progress toward nativelike production patterns may seem to appear and disappear.

My goal, then, is to examine whether progress in the acquisition of certain apparently fossilizable features, in particular, features of inflectional morphosyntax, can be detected within the operations of L2 Russian comprehension systems. First of all, that requires detecting possible changes in comprehension mechanisms over time. In addition, it would be of interest to detect systematic patterns in the L2 developments, which could yield insights into its precise nature. Thus we are led to the following two questions:

(1) Do adult learners of L2 Russian develop the ability to make use of inflectional morphology in on-line comprehension?

(2) Is there a pattern in the development of sensitivity to inflection and if so, what is its nature?

With regard to question (1), it should be noted that some researchers have suggested that many aspects of grammatical learning might be avoided in the process of learning to comprehend the L2. Gass (1997) (following Swain 1985) argues that producing spoken output forces a shift from semantic processing to syntactic processing, leading in turn to syntactic learning. If syntactic knowledge (including morphosyntactic knowledge) is modality neutral, as is commonly assumed, the comprehension system ought not to be isolated from changes in that system of syntactic knowledge. Therefore, we might more naturally expect that once some element of syntax knowledge is developed through the stimulus of spoken output efforts, comprehension too should cease to be exclusively semantic and increasingly show a nativelike balance of syntactic and semantic processing. Thus, given at least a moderate degree of apparent gains in the level of nativelikeness of spoken production, a negative answer to question (1) would require some explanation. On the other hand, even if syntactic knowledge is not modality neutral (in line with much of the thrust of Chapters 1 and 2), the existence of "production syntax" in the absence of "comprehension syntax" would at the very least set L2s apart from L1s in a fundamental

way. From a functionalist linguistic perspective and from a psycholinguistic processing perspective, the elements of linguistic form play a role in comprehension processes. It is hard to conceive of L2 speakers acquiring the ability to provide their listeners with formal patterns that for them personally have no function. If they were to succeed, whatever the L2 users would have developed would not be a normal language system with the production sub-system conforming to the expectations of the comprehension sub-system. The comprehension sub-system would not even *have* the relevant expectations. Finding solid data in support of such a fundamental L2 peculiarity would be of interest. On the other hand, if a truly linguistic comprehension system is in fact acquired by L2 learners, even if only as a result of getting things right syntactically in the production system, then we would have to wonder why the comprehension system is not subsequently sensitive to the non-target-like productions that reemerge in the classic fossilization scenario.

Turning to question (2), the discovery of systematic patterning in the development of L2 inflectional processing tendencies would be of interest in that it might suggest that the acquisition of comprehension ability associated with grammatical morphology proceeds in a principled, orderly fashion. This would in turn give hope that at least in the area of inflectional morphosyntax, L2 acquisition is *language* acquisition in some reasonable sense. In any case, if we were to find evidence of non-random developmental patterns, this could contribute to our understanding of the mechanisms involved in the L2 acquisition of inflection.

Questions (1) and (2) also address what might be considered the functionalist-processing counterpart of the access to Universal Grammar (UG) debate in SLA research (Eubank, 1991; Epstein, Flynn & Motoharjono, 1996; Schwartz & Sprouse, 1996). That debate revolves around the question of whether UG (the hypothesized innate system that enables children to learn their L1) plays a role in SLA. Bley-Vroman (1989) and Schwartz and Sprouse (1996) agree that the learner's L1 system plays a major role in the L2 system at the outset, and agree that the learner's L2 system changes over time. However, Schwartz and Sprouse maintain that the ongoing developments have full access to UG. Bley-Vroman, on the other hand, maintains that the only truly linguistic aspects of the L2 are those carried over from the L1, while any further learning will involve nonlinguistic cognitive mechanisms. If a case can be made that a complex inflectional system such as that of Russian is essentially linguistic in nature, and if questions (1) and (2) are answered in the affirmative, then this would appear to favour the spirit of Schwartz and Sprouse's position over Bley-Vroman's, even though the questions are not posed here within a UG framework. Bley-Vroman treats ongoing SLA as entirely nonlinguistic (beyond the contribution of the L1). Finding apparently linguistic L2 developments could be taken as contrary evidence to such a view.

In relation to questions (1) and (2), the best news for language learners (given common, if unrealistic, social expectations that L2 users ought to become nativelike) would be that the comprehension functions of inflectional morphology are acquired, and that their acquisition is systematic. On the other hand, we would also like to shed light on the problem that Selinker raised regarding the persistence of nonnative patterns. When temporary nativelike production patterns emerge, what is behind their emergence and when they disappear, what has become of the earlier basis for their emergence? An obvious possibility, supported by research evidence (see Chapter 4 on the effects of production planning) is that short-lived nativelike production patterns reflect metalinguistic planning and not normal language mechanisms. It is possible that metalinguistic strategies would not have a great impact on comprehension processes, while enabling spoken output to appear more nativelike. We might explore this empirically by asking a third question:

(3) What, if any, is the relationship between metalinguistic (analytical, formal grammatical) strategies and inflectional processing during L2 listening comprehension?[3]

It might be natural to expect that metalinguistic strategies would play at most a limited role in listening comprehension. Thus investigating what is happening with regard to some linguistic element in comprehension could ultimately help to sort out the linguistic and metalinguistic aspects of learners' performance in tasks that allow freedom for the use of problem-solving strategies. There are few purely metalinguistic tasks. Rather, metalinguistic reflection on L2 stimuli are carried out over and above linguistic processing of the same stimuli. In L2 spoken and written production (especially in instructed contexts) and in L2 grammaticality judgment tasks involving written stimuli, there are possible contributions of both sorts of processes. Masny and d'Anglejan, (1985) attempted to determine the extent to which linguistic ability plays a role in L2 grammaticality judgments. They used intended measures of linguistic ability (including classroom grades on free written and spoken production tasks, and cloze test scores) which may themselves have included metalinguistic as well as linguistic components, just like the grammaticality judgment tasks with which they were compared. This illustrates the need to determine just what portion of the performance in some task that allows analytical reflection would also be in evidence in an analogous task that eliminated the opportunity for analytical reflection. This is exactly what will be attempted in connection with question (3).

If the answer to question (3) is that metalinguistic knowledge does not influence comprehension processes, then we might have an explanation of the reemergence of non-target-like patterns in spoken production that were believed to have been eradicated. The metalinguistic production strategies may be abandoned, or the explicit memory on which they depend may decay. Perhaps it is not until full-fledged linguistic developments have occurred, reflected in comprehension as well as production, that the target-like patterns will be stable. Alternatively, it is conceivable that metalinguistic production strategies could in fact become automatized or proceduralized, like some other non-linguistic cognitive processes. In either case, it is important to see whether various details of linguistic form (including inflectional form) are playing a role in the L2 comprehension system.

Questions (1) - (3) will be explored particularly in Chapters 3 and 4. In Chapter 3, I attempt to determine empirically the extent to which various inflectional categories are playing a role in the listening comprehension of L2 Russian users. Participants are divided into two groups, based on the time they have been learning Russian. In all, data are analyzed from ninety-one L2 Russian users and twenty-three native-speaking Russian control participants. It is thus possible to suggest possible implications (in a pseudo-longitudinal manner) regarding the nature of the changes that occur over time with respect to inflectional processing in comprehension. In addition, the particular inflections explored belong to various theoretically motivated categories. The data will also include a

---

[3] In this dissertation, the term metalinguistic is used in a narrow sense related to what is commonly called "explicit grammatical knowledge", rather than in the broadest sense of awareness of any thought, attitude or feeling about linguistic expressions (see, e.g., Tunmer & Herriman, 1984), or even the somewhat narrower sense that includes judgements of "acceptability, ambiguity, synonymy, ordinariness, meaningfulness, comprehensibility", etc. (Chaudron, 1983, p. 345). Rather, correcting a gender error "by feel" without reflection is considered here to involve linguistic mechanisms. Making a conscious evaluation of the gender of a noun, and a then making a conscious decision regarding how to form an agreeing adjective, is a metalinguistic strategy in the narrower sense intended here.

potential measure of metalinguistic knowledge (at least potentially) to compare with the evidence regarding on-line use of inflectional form during comprehension.

Before getting to that research, however, it seemed it would be helpful and perhaps even necessary, to travel a small distance afield in Chapters 1 and 2. The logic of the experiments reported in Chapter 3 depends on a particular understanding of language comprehension. In the experiments, the sentences that are presented for comprehension (and subsequently for metalinguistic analysis) contain various comprehension cues that clash with one another for native listeners. The assumption is that if the cues are functioning in the nonnative comprehension systems, then the same cue clashes will be experienced by nonnative listeners, causing similar reactions, even when there is little opportunity to consciously reflect on the form of the sentences. The idea of cue clashes depends on an understanding of language comprehension as consisting of a cascade of processes in which cues in the speech stream trigger the events ultimately ending in the construction of conceptual representations (understandings). Over the years, considerable evidence has accumulated regarding the nature of conceptual representations, the nature of acoustic cues, and the nature of various other cue systems mediating between acoustic cues and conceptual representations.

It seemed worthwhile to survey a broad sweep of such research. Beside providing the general empirical context for Chapter 3, the understanding of language processing presented in Chapter 1 hopefully opens the way to a simplified understanding of language learning. Various scholars (e.g., Sharwood-Smith, 1986; Carroll, 1999; Pienemann, 1999) have felt a need for a learning mechanism that is separate from the comprehension system, though dependent on it, and have assumed that this learning mechanism must produce structural knowledge such as is required by various theories syntax, phonology, etc. Despite theoretical progress (especially, Carroll, 1999; Pienemann, 1999) the learning device remains somewhat of a black-box. Given the understanding of comprehension, production, and syntax presented in Chapter 1 of this dissertation, some of the mystery fades and comprehension and learning become aspects of the same process. The experiments reported in Chapter 3 are not intended to directly support the conception of learning presented in Chapter 1, any more than research based on, say, the Principles and Parameters theory is intended to directly support that theory. Rather, a theory is to be valued if it enables us to ask interesting questions. The specific questions asked above and in Chapter 3 cannot be properly understood except against the background of assumptions developed in Chapter 1. In fact, Chapter 1 amounts to a partial cataloguing of support for the interpretation of grammatical form as a system of processing cues. The view of learning that is proposed in Chapter 1 arises naturally from that view of grammatical form. Chapter 3 supports this view insofar as it demonstrates that this view of learning makes certain varieties of data interesting and meaningful in new ways.

Chapter 2 presents a discussion of inflectional morphology, which is the more specific domain of the experiments reported in Chapter 3. It is an expansion of one aspect of Chapter 1. Under the rubric of *grammatical cues* in Chapter 1, there is a more specific variety called *grammatical morphemes* and a subset of those called *inflectional cues*.

Whereas the discussion in Chapter 1 is primarily psycholinguistic in nature, building mainly on experimental evidence, the discussion in Chapter 2 is more linguistic in nature, although the emphasis is still on processing. From a processing perspective, grammatical morphology is lively and active. Languages are ever in the process of developing new grammatical morphemes, a fact that further increases the sense that they play an important role, and are not just accidental details of form, redundant—unimportant and hence readily neglectable by learning mechanisms. The nature of inflectional cues, that is, the forms they take and the processes they trigger or

constrain, is particularly complex. Chapter 2 also discusses the challenge that this complexity presents to the learning mechanism proposed in Chapter 1.

Chapter 2 also introduces Russian inflectional systems, and reviews major streams of research related to inflection in SLA studies. As with Chapter 1, Chapter 2 is intended to provide context needed for a full understanding of the work reported in Chapter 3. The liveliness of inflectional cues will be seen in the performance of the native Russian control participants in Chapter 3. In the case of the experimental participants, we will want to see whether the same cues are inert, lively, or in the process of coming to life.

As noted, in Chapter 3, I attempt to detect evidence that L2 Russian users learn to exploit inflectional form in listening comprehension. I also investigate the possibility of a relationship between such processing ability and what might be considered the relevant metalinguistic (so-called explicit grammatical) knowledge. In that all of the learners are instructed learners, I was interested to see the extent of the difference between the tendency to detect errors when participants were free to analyze the sentences and the tendency to detect errors when their attention was diverted from linguistic form and strongly focussed on meaning, given errors that are highly detectable by native speakers under the latter condition.

Chapter 4 attempts to flesh out and extend the findings of Chapter 3, and to explore some of their implications for various models of SLA and for research paradigms in SLA studies and applied linguistics.

1. Acquisition of *What*?

The phrase *second language acquisition* has an appearance of straightforwardness that risks begging fundamental questions. What in fact, is acquired? A mature second language user has come to possess some internal mechanism(s) that s/he did not possess prior to his or her experience with the second language. It is common to speak of the acquired entity as a *grammar* (perhaps an *L2 grammar*, or an *interlanguage grammar*). For example, Eubank (1994) speaks of a stage in the acquisition of L2 English when the learner would be

> ...consistently observing the [third person singular agreement pattern] in the input while perceiving that it is disallowed in his or her own grammar... (Eubank, 1994, p. 91)

until certain revisions are made in the internal grammar. Similarly, Gass (1997) discusses *intake* as a process in which

> [i]nformation is matched against prior knowledge and where, in general, processing takes place against the backdrop of the existing internalized grammatical rules. (Gass, 1997, p. 5)

That is, the set of internalized rules is matched against further samples of the language, and revisions may be made if the rules as formulated turn out to be incompatible with the new samples. To take a hypothetical example, on this view, a nonnative English user might produce utterances such as *John the poster read*, based on his or her current knowledge state. Additional input will lead to a revision of this knowledge state, as it is discovered that in English, people would not say *John the poster read*, but rather, *John read the poster*. The change involves the learner going from a knowledge state which allows *John the poster read*, to one which does not allow such sentences, but rather requires sentences of the form *John read the poster*. Acquiring the L2 amounts to acquiring knowledge of what one can and cannot say if one is to produce (and avoid producing) utterances which native speakers would produce (and avoid producing). As a further example of this orientation to SLA, consider the following quote from Mitchell and Myles (1998) in which they are discussing John Anderson's ACT* model (Anderson, 1983) as it might be applied to a specific detail in the acquisition of L2 English:

> If we take the example of the third person singular *-s* marker on present tense verbs in English, the classroom learners might initially know, in the sense that they have consciously learnt the rule, that *s/he* + verb requires the addition of an *-s* to the stem of the verb. However, that same learner might not necessarily be able to consistently produce the *-s* in a conversation in real time. This is because this particular learner has declarative knowledge of that rule, but it has not yet been proceduralized. After much practice, this knowledge will hopefully become fully proceduralized and the third person *-s* will be supplied when the conversation requires it. (Mitchell & Myles, 1998, pp. 87-8)

This quote aptly illustrates the tendency to view SLA as a matter of learning how to form utterances as required by the target language (TL) norms. This assumption regarding the *what* of L2 acquisition is rarely questioned.

An alternative orientation to SLA that is less commonly observed involves viewing the elements of language form in less static, more lively terms. Each aspect of linguistic form has a role or roles to play in comprehension. The roles consist in triggering or constraining various processes involved in understanding utterances. This approach is illustrated by SLA research drawing on the Competition Model of Bates and MacWhinney (Bates & MacWhinney, 1989; see Coorman & Kilborn, 1991 and MacWhinney, 1997 for surveys of SLA-related research in this framework). From the perspective of the Competition Model, elements of linguistic form are first of all *cues* that trigger or constrain comprehension processes. What must be acquired includes the correct mental responses to specific formal cues. During comprehension, reactions to some potential but misleading cues may need to be suppressed, as in the case of lexical or syntactic ambiguities. In order for language comprehension to be successful, the cues selected within the speech stream need to be those that best converge with one another and with the comprehender's understanding of the world. Speech *production* is a matter of providing listeners with such sets of convergent cues. "Ungrammatical" L2 utterances do not result from simply not knowing how to arrange words for the sake of spoken output (or put differently, from not having a grammar which specifies the same set of sentences and structural descriptions as the native grammar specifies). Rather, they indicate that sensitivity to some cues has not developed to nativelike levels for use in comprehension processes, a situation which ripples through to the production system as discussed below.

Thus, there are at least two orientations regarding the *what* of SLA: (1) What is acquired is abstract knowledge of how the parts of utterances are arranged in TL sentences; vs. (2) What is acquired is a set of mental responses to processing cues (and secondarily, speech production strategies aimed at providing listeners with convergent processing cues). It seems likely that the choice between these two orientations is consequential, probably in absolute empirical terms, but at the very least, in terms of the varieties of data the researcher will be drawn to examine. The remainder of this chapter aims to further sketch the view of the elements of linguistic form as processing cues. This will provide the framework in which the research questions pursued in Chapter 3 will be of interest. In addition, in Chapter 4, it will be argued that the orientation to SLA as grammar acquisition is less compatible with the gradual nature of L2 development observed in Chapter 3 than is the orientation to SLA as the acquisition of processing cues and reactions to such cues.

1.1. FROM VIBRATING AIR TO UNDERSTOOD MESSAGES

Recently at an airport I heard a group of tourists talking to one another while they waited to get on a bus. They were speaking in a language I didn't recognize, much less understand. In such situations, about all one is aware of is that the language is a language. Beyond that, one may be aware of some auditory characteristics of the sound stream. The subjective experience of those people as they use their language was on an entirely different order from my subjective experience of them using it. Typically, the sound stream as such would be out of their awareness. They have the illusion as they listen to one another that they are directly hearing one another's ideas. The same sound stream that is an opaque wall to me would seem to be a transparent window for them. This illusion is remarkable, considering the enormous amount of computation that is taking place in their heads and the speed at which the events of normal speech occur. There is

something inside the native speakers and missing in me which enables them and not me, to experience this flow of sound immediately as if it were a flow of "ideas". It was not always that way for them, either. They were born into an environment that had this language as a major feature. The language which surrounded them was there preceding them and will be there after them. What had to develop in them was a language processing system that would take the speech sounds in their environment and convert them into "ideas" at high speed. And now, the contrast between my experience of their language and their experience of their language points to the existence of something truly phenomenal that is present in them and absent from me.
If I were to follow those tourists around for several days (perhaps even if I didn't pay a whole lot of attention to them) and then soon afterward were to hear another group of tourists speaking the same language, I would likely recognize the language as the one I heard before. The fact that I recognized the language would be proof that I had learned something. That is, properties of the sound stream are now able to trigger recognition in me. In some sense, my learning of this language would have begun. It might not go much farther. On the other hand, under the right circumstances my learning might continue until my subjective experience of those people's language might be similar to theirs in many ways. Eventually, again given the right circumstances, their language could even come to be the primary language with which I deal in my life. From a native-speaker perspective, my L2 system might well be mildly to moderately dysfunctional (Coppieters, 1988, Johnson & Newport, 1989). Nevertheless, adult-acquired language processing systems, when the L2 linguistic history of the adult has been adequately rich, achieve most of what native language processing systems achieve, and thus merit comparable awe.

Native language processing systems have been studied on an enormously greater scale than have "near native" ones. However, in the broadest sense, these two varieties of language processing systems solve the same overall computational problem: they create understanding from sound. Listeners allow the sound flow to constrain the nature of their thoughts in systematic ways. Given the rapidity of speech, we might want to say that properties of the sound stream *trigger* events in the conceptual stream. We will see that this *triggering*, rapid as it is, must in fact involve a chain of triggerings with temporal dependencies on one another: sound must be dealt with before words can be identified, before meanings can be retrieved and so on.

1.2. THE END-PRODUCT OF LANGUAGE UNDERSTANDING

The view of language learning developed below involves associating patterns in the speech stream with patterns in the stream of understanding. An understanding of the nature of both streams would appear to bear on the question of what is being acquired. The conception of inflectional morphology that is discussed briefly in this chapter and expanded upon in Chapter 2 depends upon recognition of the reality of the events involved in constructing and revising the stream of understanding.

1.2.1. Demystifying the "stream of meaning"

The process converting sound into meaning operates on a physically instantiated sound-stream and leads to a physically instantiated conceptual stream. The instrumental means for analyzing the rapid and complex fluctuation of air pressure that constitutes the sound-stream may give us the sense that it is more tangible than the conceptual stream. In fact, the auditorily perceived speech stream and the conceptual stream must both ultimately be tied to the way in which the world is perceptually experienced and the way

perceptual experience is remembered. A stream of perceptual experience can be stored and retrieved (or reconstructed) in its own right. Someone may observe a sequence of events and then verbalize these remembered events for the first time several hours later. Apparently, the speaker has carried around something in his or her head that constitutes the memory of those events. S/he is later able to "play back" that memory at will, and aspects of the played-back memory trigger the articulatory actions involved in relating the event verbally. The sound-flow thus created then triggers *understanding* in the listeners. That is, the listeners are able to imagine the events being narrated. The ability to comprehend narrative is the ability to *imagine*, constrained and guided by the auditory stream. The ability to imagine is the ability to build "memories" of events that were not experienced (and may not have even occurred). The building blocks are the components of experience, as presented by perception and emotion and as stored and retrieved (or reconstructed). Thus although we cannot record this stream of ideas in a way analogous to recording the sound stream with a speech spectrograph, we must not doubt its reality, and we can employ some reasonable, if still hazy, hypotheses regarding its nature. The ability to sense and experience gives rise to the ability to imagine, which in turn give rise to the ability to comprehend speech (Paivio, 1971; Johnson-Laird, 1983; Kosslyn, 1994). I will use terms such as *concept*, *conceptual flow* and even *meaning* to refer to this non-linguistic level of mental representation of the world. (There is a risk in this of losing sight of the crucial social dimension of meaning. See, e.g., Putnam, 1975. If concepts are rooted in experience, their role in language is rooted in *shared* experience.)

On an informal, anecdotal level, the separation of form and meaning is a familiar phenomenon, as in the experience of word-finding difficulties. Recently I heard the Russian word *zholch*, in a context such as, "the liver starts producing *zholch*" and immediately understood which substance (based on my world knowledge) *zholch* referred to, although only with some difficulty was I able to come up with the English word *bile*, which refers to the same substance. Such clear cases of knowing "what it is" and having difficulty with "what it is called" clearly demonstrate the separation of form and meaning in relation to individual words.

According to Paradis (1994), a speaker is generally unable to repeat his or her own words verbatim if the utterance exceeds something like twenty-five words, although the speaker can reverbalize "the message" without difficulty, using slightly different words. In an influential article, Sachs (1967) demonstrated empirically and rigorously this independence of the level of meaning from the level of linguistic form. Participants who heard *He sent a letter about it to Galileo, the great Italian scientist*, if questioned immediately afterward, could recognize purely formal changes such as that in *A letter about it was sent to Galileo, the great Italian scientist*, or *He sent Galileo, the great Italian scientist, a letter about it.* However after only 80 syllables, the listeners could not correctly identify which of these sentences they had heard, although they knew they had not heard the sentence *Galileo, the great Italian scientist, sent him a letter about it.* In other words, they had accurate memory for "what happened", that is, for the meaning of the text, but not for the form which had triggered the construction of the meaning representation.

Jarvella (1971) provided evidence that the preservation of meaning and disposal of form which Sachs observed occurs particularly after sentence boundaries. What appears to happen at the sentence boundary (and to a lesser extent at clause and phrase boundaries) is that the linguistic form is (often) largely discarded, having done its job. *Understanding* appears to be realized in a format that is more readily retained in long term memory than is linguistic form.

1.2.2. Mental models and language understanding

While the evidence uncovered by Sachs and Jarvella pointed to the reality and independence of both the level of linguistic form and the conceptual level, Bransford, Barclay and Franks (1972) provided more detail regarding the constitution of the conceptual representations, presenting evidence that the conceptual representations of concrete descriptive sentences are depictive (as opposed to propositional). That is, what is remembered after hearing a verbal scene description is (in some sense) the scene itself. In their experiment participants heard, for example, the sentence *Three turtles rested on a floating log and a fish swam beneath them.* Subsequently they might mistakenly recognize *Three turtles rested on a floating log and a fish swam beneath it* as being what they had heard, but were much less likely to mistakenly recognize *Three turtles rested beside a floating log and a fish swam beneath them.* In other words they would tend to accept sentences compatible with the situation depicted by the original sentence, i.e., with the same items in the same spatial arrangement, but tend not to accept sentences whose understanding required a different depiction.

The depictive mental representations suggested by Bransford, et al. became the basis for the development of a line of research into the *mental models* that are now widely taken to be the end result of discourse comprehension (see Wilson et al., 1993). One of the major proponents of this viewpoint, Phillip Johnson-Laird (1983, 1989; Mani & Johnson-Laird, 1982; Johnson-Laird & Garnham, 1981) argued that mental models have broader utility than their use in language comprehension, in that their construction and manipulation is used in reasoning and problems solving.

An important fact about mental models is that they develop in real time. That is, they must be constructed, which implies that in some sense they are built out of simpler components. It appears that construction (in the sense of building complex forms from simpler components) is a characteristic of human information processing. For example, in visual perception, information regarding colour, shape and movement are processed separately and must be integrated into unified percepts of coloured, shaped, moving objects. Visual imagery (imagining) also appears to involve a constructive process in which parts are put into spatial arrangements and so on (Kosslyn, 1994). If the conceptual processes involved in language comprehension are derived ultimately from perception, then we might expect conception in general to be constructive. MacWhinney (1997) suggests that concept construction during language comprehension overcomes memory limitations:

> The formation of the link between *black* and *dog* involves more than the simple positional relation of two words. Instead, the adjective is applied to the noun on the conceptual level and an image of a black dog is activated. The assumption is that, as soon as verbal material can successfully access an integrated conceptual representation, it no longer exacts additional storage cost. (MacWhinney, 1997, p. 133)

The appeal of this hypothesis of concept composition will increase as we see the broader picture of the relationship of grammatical composition to concept composition. Grammatical morphology, it will be argued, is related not to the components out of which the conceptual stream is built, but rather to the events of the concept building process (and certain other features of the concepts, such as their truth status).

In constructing and modifying a mental model, various humans, animals, or physical objects are added, or restored, or simply maintained in a highly active state. These are commonly referred to as *tokens* in the mental model, corresponding to what

linguists might call (mental) *referents* or *participants*. For example, Glenberg, Meyer and Lindem (1987) presented readers with one of two versions of a text, one in which a sweatshirt was put on by the protagonist and the other in which it was removed. Evidence was found that the "sweatshirt token" was kept active in the mental model of the former text, but not in the latter.

Another important feature of tokens in mental models is the fact that they serve as "pointers to collections of information in long-term store" (Glenberg, Kruley and Langston, 1994; Sanford and Garrod, 1981; Garnham, 1981; Givón, 1990). Listeners appear to add information about a character over time while not remembering where in a narrative the information was added (Garnham, 1981). Givón (1990) refers to referential descriptions as pointers to *files* in episodic memory. As the mental model token participates in various events and scenes, material is added to such a file and can later be used to identify a referent, or assumed for inferential purposes.

Although mental models are rooted in the way the world is perceived, recalled and imagined, it is also the case that languages add a new dimension to memories of mental models (as in fact they add a new dimension to direct experience) in that there is the possibility of remembering aspects of the verbal reports as well as the mental models derived from the verbal reports. The remembering of text that does not readily resolve into a single determinate mental model is made possible by the ability to remember aspects of the verbal form (Mani and Johnson Laird, 1982). Also there are various other conditions which favour the retention of memory for verbatim verbal form (Kintsch & Bates, 1977; Keenan, MacWhinney & Mayhew, 1984).

Abstract language adds another dimension. What is a mental model of the noun phrase *factors contributing to the retention of exact verbal form*? We can conjecture that models based on abstract discourse are dependent on the format of concrete models. That is, in the phrase at hand, there is a metaphorical token in the model for *factors*, as though *factors* are physical objects, and there is a file (collection of properties) associated with that token. In fact, metaphor is a common characteristic of language, contributing to abstractness (Lakoff & Johnson, 1980). The success of metaphors depends on the success of the analogies they invoke. *Blakonia, provoked her neighboring countries by repeated small but irritating provocations in the form of territorial incursions, until the patience of all the neighboring countries wore thin and they ganged up against Blakonia and crippled her military capacity for the foreseeable future*. This one-sentence text might invoke a mental model in which Blakonia, an imaginary nation, is represented by a token analogous to a token that would represent a human protagonist in a concrete narrative, and the neighboring countries are also treated as human neighbors who can be irritated, and if irritated enough, will act on those irritations against the one causing them. (Some abstract language may not get beyond such metaphorical mental models, as in the case of mathematical or scientific discourse.)

Although the presence of language may, in important ways, give special characteristics to mental models derived from verbal sources, we should not let this cloud the importance of the ordinary concrete mental model as a foundation for the system of language understanding. The following anecdote from my own recent experience is intended to put our feet back on the referential ground as we think of the comprehension of concrete narrative. I was listening to a tape of a Russian text (based on an English text in Romjin & Seely, 1988). It consisted of a set of instructions for getting into a car, starting the motor and driving away. In listening to the text, I heard an instruction to open the car with the key, where the Russian verb *otkroj*, 'open:familiar.imperative' occurred with the direct object *mashinu*, 'car:accusative'. This was followed by an instruction to open the door, employing the same verb and the direct object *dvertsu*, 'door:accusative'. At that point, my comprehension system ran into major trouble. In my understanding, it

was impossible to open the door, because the door was already open. That is, following the event of opening the car with a key, my narrative understanding had a "mental car" with its "mental door" open, and the new instruction to open the door clashed with this mental state of affairs. I was forced to reinterpret the original opening of the car in terms of simply unlocking the car, and then the further processing of the discourse became possible. When I talk about a mental car with its door open, I mean that I am capable of mentally representing cars and car doors in ways closely connected to my sensory experience of cars and car doors, and I am able to represent acts of unlocking and opening also in a way closely connected to non-linguistic experience (perhaps motor-kinesic experience). In admitting that I misunderstood the verb *otkroj* on its first occurrence, I am accepting that there is such a thing as "understanding" that is non-linguistic in form and tied to representations of things in the world. The available evidence suggests that my mental model contained activated units corresponding to objects in spatial relationships, associated property units, and events such as motions.

We now have three varieties of streams (i.e., three varieties of changing patterns of mental activation). There is the general stream of sensory experience, including the subpart of it that is selected for attention and remembering. Secondly, there are similar streams that do not represent immediate experience. They are either based on memory of previous direct experience (and hence in some sense *re*constructed) or constructed de novo using the same building blocks (as with inferred or imagined events and event sequences). Thirdly, we have the auditory stream of speech. The stream of immediate experience, if not the stream of constructed representations of experience, is for all intents and purposes continuous, as is the stream of sound. The capacity of the one stream to trigger the creation of the other (going in either direction) depends heavily on the ability of the human organism to categorize and unitize. Categorization provides a manageable set of building blocks at both ends (units within the sound stream and units within the streams of experience and understanding). The categories used for mental model construction include categories such as tables and chairs, running and walking, redness and greenness and so on. The categories needed for processing the sound stream include speech sounds, phonological words and so on. Language learning will thus include category learning in relation to the stream of sound and the stream of experience/understanding.

1.3. USING CUES TO CONSTRUCT UNDERSTANDINGS

Having somewhat clarified the nature of understandings, we are in a better position to address the question of what is acquired when a language is acquired. The learner must learn to construct mental models in response to cues in the speech stream. Certain cues in the speech stream will trigger the addition of tokens to mental models. Other cues will trigger (or constrain) the merger of properties (e.g., colours, sizes) with tokens, will trigger the construction of representations of events or states of affairs in which those tokens participate, and will trigger certain other operations. These include as the assignment of epistemic status (e.g., level of certainty) to the situations and events in the mental model, and the staging of situations and events in relation to one another (to be discussed especially in Chapter 2). Having given some substance to the notion of the end-point of language understanding, we turn to the speech stream and begin to examine the cascade of cues that lead to the ultimate construction of mental models.

As we now work our way back from acoustic form to mental models, we will see that the units of processing, that is, the cues, are not simply received, but rather created from evidence. The entire cue-interpreting system ultimately attempts to construct mental models which are consistent with the entire set of cues at every level (sounds, words, word

combinations), which are consistent with the comprehender's current discourse model (based on what s/he has heard up to that point) and which are consistent with his or her understanding of the world.

1.3.1. Phonetic cues

From the learner's perspective there are two starting points: an auditory stream and a stream of understanding. The auditory system in turn arises from an acoustic stream—rapid fluctuations of air pressure interacting with the elasticity of the ear drum, which in turn triggers a chain of further events. For completeness (and also because it provides a useful illustration of the concept of processing cues), we need to turn briefly to the topic of the acoustic stream and its relationship to the listener's categorization and recognition system for speech sounds. The acoustic level provides the initial cues in the cue cascade on the way to mental models. The would-be learner of a given language faces his or her first hurdle at this level. Difficulties at this stage will affect all that follows. It is also possible that if this stage were especially processor-intensive for L2 users, it might decrease the processing capacity available for other stages.

If we examine the acoustic stream as reflected in a speech spectrogram, we will be able to distinguish certain changes in patterning over time. For example, we will notice an alternation between pulsating (periodic) stretches (reflecting vibrations of the vocal cords), silent sections and sections where there is non-periodic noise. Larger portions of the pulsating segments will also display concentrations of energy (formants) around certain frequency levels, characteristic of vowels. Thus segmenting the vowels and consonants, allowing for fuzzy borders, may appear to be reasonably straightforward in principle, even if somewhat complicated in the technical details. Beyond the crude segmentation into segments and/or syllables, the identification of specific phonological segments becomes extremely complex.

Consider, for example, the categorization of an exemplar of the English voiceless palato-velar stop /k/. In the spectrogram there is an absolutely silent spot corresponding to the point in time when we might consider the mouth to be in the formation required for this consonant. Since /p/ and /t/ are also silent, more information is needed if the specific consonant is to be identified. The listener is dependent on a set of acoustic cues (Liberman, Delattre & Cooper, 1952). These include a delay in the onset of voicing of about 60 milliseconds, a cue which distinguishes the /k/ from its voiced counterpart /g/ (after which the onset of voicing would occur much more quickly, perhaps 10 ms after the onset of the burst). The burst of noise itself contains frequencies within a certain range. However, the burst frequency only helps to identify the consonant when it is combined with a vowel. That is, a frequency burst centered around 1500 Hz will be perceived as a [p] before the vowel [i] or [o], but as a [k] before the vowel [a]. More important than the noise burst, however, is the formant transition from the [p] to the vowel, which is the strongest cue to the identity of the place of articulation. This means that the primary cue that identifies the /k/ in /ka/ is physically a quite different cue from the primary cue that identifies the /k/ in /ki/ (Delattre, Liberman & Cooper, 1955). This gives just a small indication of the complexity involved in identifying "speech sounds". The intuition that the sounds are simply strung together, floating along in the stream one after another ready to be recognized, is misleading. Rather, the mental language processor receives separate bits of evidence, some of which are temporally somewhat distant from the perceived site of the phonetic segment. From these bits of evidence, the processor computes the identity of the sounds, creating whatever sense we may have of separate phonetic segments. Adult L2 learners have an L1 speech processor which is extremely good at ignoring irrelevant aspects of the sound stream and reacting decisively to relevant ones.

Studies of L2 speech perception have indicated that the new cue system is not readily acquired, and that the L1 cue system affects the way acoustic cues are used in processing the L2. Native English listeners hear at least two separate cues that help them to distinguish the words *peace* and *peas*. The vowel of *peace* is shorter than the vowel of *peas* and the final consonant of *peace* is shorter than the final consonant of *peas*. (In neither word is the final consonant voiced.) Flege (1991) reports that the consonant length cue had no effect on L2 English listeners from Finnish and Swedish L1 backgrounds. Vowel length served to distinguish the English word *peace* from the word *peas* for these same L2 listeners, but Flege argues that they may have been using the vowel length cue as a contrastive property of the vowel, in accordance with their L1 phonetic cue systems, and not as a cue to the identity of the consonant, as in the target English system. Cutler, Mehler, Norris and Segui (1992) found that in certain respects highly proficient French-English bilinguals used only the cue system of their self-declared dominant language and only in processing that language.

On the other hand, it appears that learning does go on in this domain of L2 speech processing. In my language learning diary, I commonly noted for over a year that my ability to hear Russian pronunciation (that is, the degree to which I perceived it as different from English) continued to improve. If this observation was valid, then some variety of auditory phonetic learning must have been going on over all of that time. Leather and James (1996) cite MacKain, Best and Strange (1981) as having presented evidence that L1 Japanese learners of L2 English with high proficiency displayed relatively nativelike use of the acoustic cues distinguishing English /l/ from /r/, while learners of lower proficiency did not.[4]

This is interesting in connection with the picture that emerges in Chapters 3 and 4 of this dissertation regarding the time-frame of inflectional acquisition in L2 Russian. It is argued there that this aspect of second language development involves the gradual strengthening of cues over a number of years. The findings of MacKain et al. (1981) open the possibility that this can be true in the domain of L2 speech perception as well.

The ability to make use of phonetic cues must directly impact the issues of inflectional processing in Russian. For example, the suffix -e marks two different noun cases in Russian (one of them across two declensions), making it a relatively frequent case ending. In theory at least, this case ending could often be identifiable to a native Russian listener during the preceding syllable. That is because the consonants preceding the (front) vowel of this suffix are palatalized, and for native Russian listeners the formant transitions preceding the consonant would be a clear cue to the palatalization of the consonant (Bondarko, 1998). In Experiment 1 of Chapter 3, there is a crucial contrast between the nominative case form *kreslo*, 'chair', and the locative case form *kresle*. To a Russian listener, there are separate, clear cues to the case form spread across the portion *...esle* (or *...eslo*). For L2 Russian listeners of many language backgrounds, if they are dependent on their L1 cue systems, the contrast between the two case forms may well lie only in the final vowel. In a purely phonetic sense then, the case-marking contrast is set up to have a much weaker phonetic basis for a listener employing a nonnative acoustic cue system than for a listener using a native Russian acoustic cue system. Depending on how slowly the relevant phonetic learning occurs (and from my subjective experience, I would feel that at least some of the relevant learning *can* occur), the nonnative disadvantage could be prolonged. Moreover, the absence of nativelike acoustic cue processing mechanisms could well be more general than this single example illustrates.

---

[4] The L2 Japanese participants reported on by Brown (2000) were unable to distinguish English /l/ and /r/ with greater than chance accuracy. Those participants were students in an English Canadian university, suggesting that they were at a reasonably advanced proficiency level.

As already noted, an important feature of speech perception, as of speech processing generally, is the speed at which it takes place. Kess (1992) cites Liberman (1970) as indicating that speech is processed at a rate of twenty-five to thirty phonetic segments per second, while Tartter (1998) cites Liberman et al. (1967) as providing the estimate of 900 segments per minute (i.e., 15 per second). Non-speech sounds cannot be recognized unless presented at a rate less than four per second (Kess, 1992). It is true that the cues to the identity of phonetic segments are spread out over a distance larger than the segment. However, it could be argued that this makes the task even more complex, since acoustic cues related to more than one segment are being processed simultaneously. In any case, there are a variety of indications that the stream of speech sound is processed by a system in part separate from the system that processes non-speech sounds. As Epstein, Flynn and Martohardjono (1996) point out, the fact that L2 speech perception is possible at all strongly implicates the language faculty in L2 functioning (in case there were any question).

1.3.2. Lexical cues

As acoustic cues trigger the processes involved in the categorization of speech sounds, so the speech sounds (if not the finer details themselves) serve as cues to larger, recurrent stretches of sound which in turn activate the conceptual building blocks for mental model construction. In considering the auditory lexicon, we again find a general picture of extremely rapid processing in the context of considerable complexity. It is not only that the phonetic form of each word must be matched to some representation of the word in memory in order for the identity of the word to be determined. The meaning of the word must also be found in memory and exploited (perhaps along with other aspects of the word's form) for the sake of deeper aspects of the comprehension process. The problem of lexical access is sometimes compared with the problem of finding individual words in a dictionary containing tens of thousands of words, and finding them at a rate of two and a half entries per second (in the case of normal spoken English—Marslen-Wilson & Tyler, 1980) and considerably more quickly at times.

Recognition of the phonetic forms of incoming words requires that there be some representation of those phonetic forms in memory against which the incoming stretches of sound can be matched. One question this raises is what happens when two words have the same phonetic form, as in the case of *bank* meaning the edge of a river and *bank* meaning the place where I have my checking account. Swinney (1979) provided experimental evidence that the single phonetic form shared by different homophones briefly triggers the activation of both. The activation of the contextually inappropriate meaning rapidly declines and the listener is typically not aware of it having occurred. However, when combined with Marslen-Wilson's (1987) *Cohort Model* of lexical access, Swinney's finding had an important bearing on the issue of how the auditory contact forms of words are activated during listening. The Cohort Model deals with the time course of the activation of the auditory contact forms of words. Some early findings indicated that listeners often recognize a word before they have heard it in its entirety (Marslen-Wilson, 1975; Marslen-Wilson & Welsh, 1978), often within 200 to 250 milliseconds after its onset (when the mean length of the words was 370 milliseconds). This raises the question of what happens when there are several words which are identical in their initial portions. The Cohort Model holds that at each point in time as a word is being progressively perceived (presumably after some essential minimum), all words compatible with the portion of the current word encountered up to that point in time are being activated. Zwitserlood (1989) extended Swinney's (1979) findings that the various meanings of homophonous words are activated during listening comprehension to the

investigation of groups of words that overlap in some initial portion of their phonetic form. For example, the words *comprehend* and *compromise* are compatible up to the *r* segment and perhaps even to the reduced vowel following the *r*; but by the time the *h* segment is identified, *compromise* is ruled out as a possibility. Zwitserlood found that until this *recognition point* in a word is reached by the listener, the meanings of all words with a particular initial phonetic portion showed evidence of activation, although once the identity of a word had been finally determined, the meanings related to those other words in the cohort were no longer accessed. Shillock (1990) further extended this line of research to non-initial word portions, such as the second syllable of *trombone*. Here too the evidence suggested that the meaning of the word *bone* is activated when *trombone* is heard.

Thus it appears that the auditory contact representations of lexical items, besides being activated extremely rapidly, are also extremely sensitive to the form of the input. The processor first finds possible word forms wherever they occur, while quickly discarding those that are contextually inappropriate. For English, given the evidence that not just forms, but meanings are activated and rapidly decay, we are encouraged to believe that both the process of word recognition and the process of activation of the associated meanings happen extremely rapidly. If the correct word is typically identified and understood in 200 to 250 milliseconds, the words in the initial cohorts are apparently being activated and deactivated even more quickly than that, including both their auditory and semantic aspects. To use the dictionary metaphor again, imagine an electronic dictionary in which words actually light up when activated. As the auditory stream races by, every possible word embedded in the sound stream almost instantly lights up in the dictionary with little time lag following its occurrence, and then inappropriate choices quickly fade away, while the appropriate choices become much brighter before also eventually fading. Added to all of this, yet another impressive aspect of this mental lexical system is the way that new L1 vocabulary items can apparently be added often after a brief encounter and immediately begin functioning reasonably well (see Gupta & MacWhinney, 1997).

Somewhat related to SLA is research into the nature of the bilingual mental lexicon. Much recent work in this area has been inspired by Potter, So, von Eckhardt & Feldman (1984). Potter et al. argued from experimental evidence that words which are translation equivalents in the two languages of a bilingual individual are connected to one another in the mental lexicon indirectly, via the conceptual form (meaning) that the two words share, and are not connected directly (L2 word form to L1 word form). For example, in a Spanish-English bilingual, the word forms *dog* and *perro* would not be directly connected to one another, but since both the Spanish word form and the English word form would be connected to the concept 'dog', the two word forms would be connected to one another via that concept. In related research, Kroll and Curley (1988) found a difference between L2 users with low and high proficiency (in effect, early learners versus functional bilinguals). The latter performed similarly to the participants in the Potter et al. study, while the participants who were at an earlier stage in their learning performed in a way that Kroll and Curley interpreted as implying the existence of direct links between the word forms of the two languages in the mental lexicon, so that *dog* and *perro* would be connected at the level of their forms and not just via a shared concept (similar findings were reported by Chen and Leung, 1989; see also Blekher, 1999). More importantly, these scholars interpret their results as suggesting that for low-proficiency L2 users, the mental path from the L2 word to the concept (e.g., from the Spanish *perro* to the 'dog' concept) is via the corresponding L1 word form (English *dog* in this case). In the words of Kroll & de Groot (1997) "[the] L1 initially holds privileged access to meaning" (p. 178). Over time, on this view, "direct conceptual links are also

acquired...[while] the lexical connections [that is, connections between the forms of words] do not disappear..." (Kroll & de Groot, 1997, p. 179).

On the level of form, this research tends to treat the mental lexicon in a unitary manner. That is, differences between the auditory, visual-orthographic and motor-articulatory lexicons are often not taken into consideration. Consider a mental apparatus that contains a particular L2 auditory contact representation and thus reacts to a word in the sound-stream, but in order to activate the associated concept, it must first activate the L1 auditory contact representation (or the visual-orthographic or motor-articulatory representation, or some combination) before the necessary concept can be activated and the word understood. Whatever such a mental apparatus might be, one might wonder whether it would merit the label of mental lexical entry. We might be justified in wondering whether such a system, in which L2 auditory contact representations function to activate L1 forms rather than directly activating concepts, could function under the time constraints of normal speech comprehension.[5]

In the case of the spoken production lexicon, there might be more scope for the use of what could amount to non-linguistic knowledge of L2-L1 surface word correspondences during production planning, since as speaker, the L2 user may have more control over content and form than s/he has as listener. This theme will re-emerge especially in Chapter 4 where a similar proposal will be made in connection with L2 inflections.

1.3.3. How far do lexical cues (content words) take us?

In our journey from acoustic details to mental models, this is a good point for stock-taking. The products of speech recognition to the level of sound segments or other phonological units (such as distinctive features, syllables and prosodic groupings) are of little immediate help in guiding the process of mental model construction. However, as soon as *words* are recognized by the mental language processor, the picture changes dramatically. To what extent might a listener be able to construct an adequate mental model on the basis of lexical information without the aid of further levels of grammatical information? In other words, how much does the listener get simply as part of the "meaning of a word"? The above cited research by people such as Swinney (1979), Zwitserlood (1989) and Shillock (1990) uses a technique called semantic priming to detect the activation of word meanings that the listeners are unaware of. For example, in the case of *trombone*, the participants are presented with the word *rib* and must decide whether it is a word or not. Following the word *bone*, people can identify rib as a word more quickly than they can in a semantically unrelated context, such as following the word *book*. Immediately after the word *trombone*, recognition of *rib* is facilitated as much

[5] Lotto and de Groot (1998), in one of their experiments, appear to have been able to inhibit the tendency of learners to associate L2 word forms with L1 word forms, a result which they judged undesirable, since participants who formed strong links from L1 words to L2 words were moderately more able to recall L2 words in response to cues (pictures or L1 words) than were participants who did not form such L1-word-L2-word links. They do not consider the possibility that the independence of L2 lexical entries from L1 translation equivalents might in its own right be a desirable feature of L2 lexical development. One might detect here a divergence between a pedagogical perspective on vocabulary learning and a long-term perspective on lexical development. As Jiang (2000) observes, "In first language development, the task of vocabulary acquisition is to understand and acquire the meaning as well as the other properties of the word. In tutored L2 acquisition, the task of vocabulary acquisition is primarily to remember the word" (Jiang, 2000, p. 50). Looking at findings such as those reported by Lotto and de Groot, how does one weigh the value of some particular number of concept-mediated L2 lexical entries against the value of some other number of L2-L1 word-associated vocabulary items?

as it is following the word *bone* itself. The semantic networks of specific words are thus extremely sensitive to the occurrence of the acoustic forms of those words in the speech stream. Such activation of whole semantic networks could play a role in making available rich bodies of material for mental model construction.

Not only may words activate semantically related words in a general semantic network. Meaning representations of individual words can be compared to knowing a picture with some missing parts and having some expectations regarding the parts that are needed to complete the pictures. Langacker (1991) develops this idea within a linguistic framework in terms of words *elaborating* one another's meanings. For example, the verb *cook* has a schematic meaning that can be filled in many specific ways. The word *vegetables* can elaborate that schematic meaning, making it more specific, as can the word *meat* (or even more specifically, the compound *rump roast*). That is, the ultimate form of the cooking event (or activity) in the mental model will depend on what is being cooked. The word *vegetables* is also schematic and can be elaborated in different ways, as when someone *buys*, *washes*, *chops*, *freezes*, or *cooks* vegetables. Thus in combining *cook* and *vegetables*, one is taking advantage of the way both of the schematic meanings are conventionally disposed to take on more specific forms.

One area in which this property of word meanings has received attention within psycholinguistics is in the study of what have come to be called *thematic roles*, following Chomsky (1981) (see, e.g., Carlson & Tanenhaus, 1988; Stowe, 1989; Trueswell, Tanenhaus & Garnsey, 1994). Fillmore (1968) pointed out the need for a more conceptual, as opposed to grammatical, level of representation of verb valence. For example, in the sentence *They loaded hay on the truck*, the word *hay* is the direct object, while in *They loaded the truck with hay*, the phrase *the truck* is the direct object. Yet in either sentence, that activity of loading requires some physical object or substance that moves onto some surface. Chomsky (1972), in response to Fillmore, suggested that each lexical entry would contain information regarding how the arguments of a word would be semantically interpreted in terms of such semantic roles. Chomksy's (1965) Standard Theory lacked thematic roles, but it included selectional restrictions, such as the restriction that grammatical subjects of the verb *admire* had to make reference to animate beings. The difference between semantic roles and selectional restrictions may be just a matter of how specifically words are believed to constrain their possible arguments. Does the lexicon indicate that the complement of *eat* is a semantic patient, or that it is food? Recently McRae and collaborators (McRae Spivey-Knowlton & Tanenhaus, 1998; McRae, Ferretti and Amyote, 1997) have been suggesting that thematic roles are in fact more like selectional restrictions than is generally recognized. Instead of the verb *kick* requiring a patient as its grammatical object, it creates an expectation of a plausible *kickee* as object.

If the conceptual representation of individual words includes properties such of this more specific variety of valency (the adjective *smooth*, for example, requiring a surface to flesh out its conceptual form, and perhaps even preferring a noun from a specific set of relatively frequent possibilities), then it would seem that lexical meanings provide building blocks for mental model construction *and also* provide information regarding how to combine component blocks into complex concepts. Thus in accessing the lexical items in an utterance, a listener might be carried a long way toward the construction of a mental model corresponding to the utterance.

The prospects for constructing a mental model from lexical content alone are closely tied to the use of preexisting knowledge and inferential processes. Even at the level of speech perception, background knowledge appears to influence what listeners believe they hear (Warren, 1970; Warren & Sherman, 1974). Background knowledge is further implicated in the selection of competing homonyms discussed above. In other

words, language understanding appears to involve selection from among competing candidate units at many levels (candidate phonological segments, candidate words, candidate syntactic structures, candidate conceptual representations) such that all information sources (from acoustic cues to encyclopedic knowledge) converge, that is, are found to be compatible. However, once enough evidence has accumulated, some information sources (phonetic, lexical, grammatical) may get neglected. For example, Erikson and Mattson (1981) found that listeners often do not process aspects of the spoken input that are highly predictable. Thus when asked *How many animals of each kind did Moses take into the ark?*, listeners behave as though they had heard the sentence *How many animals of each kind did Noah take into the ark?* In other words, they somehow associate the phonetic form of *Moses* with the conceptual representation belonging to the word *Noah*. Johnson-Laird (1981, cited in Garrod & Sanford, 1994) found that listeners understood the sentence *This book fills a much needed gap* to mean that the book was needed to fill a gap, while Wason and Reich (1979, cited in Garrod & Sanford, 1994) found that the sentence *No head injury is too trivial to be ignored* was readily taken to mean that head injuries should never be considered trivial and thus ignored. Such examples (see Garrod &, Sanford 1994 for a fuller survey) seem to indicate that even native speakers are capable of working (indeed in some cases even prone to work) from the lexical content to a plausible conceptual representations without the aid of structural relationships of the sort believed to be involved in grammatical processing.

In short, in relation to our question, how far can lexical content alone take us toward the successful construction of understandings of utterances (conceptual representations, mental models), the answer appears to be that it can take us very far indeed, especially when aided by the inferential processes which also occur in normal language processing.

Such facts regarding normal language processing would appear to have implications for L2 processing. Consider the "Moses illusion" discussed above. If L1 users are able to identify Noah as referent on the basis of other lexical items (*animals*, *ark*), plus encyclopedic knowledge, then for L2 users to comprehend sentences with unfamiliar vocabulary on a similar basis would not take L2 processing out of the realm of normal language processing. Only the relative mix of processes might differ between L1 processing and (early) L2 processing, there being a proportionately larger dependence on lexical information, encyclopedic knowledge and inferences in (early) L2 processing than in L1 processing and a lower dependence on specific lexical items (unfamiliar ones) and some aspects of grammatical processing in L2 processing.

1.3.4. Grammatical cues

Grammatical cues include word order cues and grammatical morphemes (in particular, function words and inflectional modulation of content words). This grouping of grammatical cues into a single set of phenomena is based on their complementary relationship to lexical cues. Grammatical morphemes, like phonetic cues and lexical cues, play a major role in language processing and we must suspect that this is beneficial, given the fact that L1 learners, at least, do readily come to make use of them, and also given the fact that they are very widespread in the languages of the world, with new ones continuously arising in languages as earlier ones erode (Bybee, Pagliuca & Perkins, 1994). Keeping in mind the demands of high speed language processing, we can imagine that the addition of grammatical cues aids the rapid reduction of indeterminacy as lexical concepts are being deployed in the construction of mental models. A natural consequence of approaching language in the manner adopted here is to view elements of grammatical

forms as yet another set of processing cues. The processor is attempting to create conceptual representations using all the evidence available and grammatical form is one important class of evidence. Such a conception of grammar is not unknown in linguistics. For example, Givón (1990) proposes to

> ...reinterpret grammar as mental processing instructions. The grammatical signals (morphemes, syntactic constructions) used to code referential coherence in discourse are designed to trigger specific mental operations in the mind of the **speech receiver** ('decoder' 'hearer'). (Givón, 1990, pp. 894-5, emphasis original)

From a processing standpoint, it could ultimately turn out that "grammatical signals" are not in fact a unified set of phenomena. However, at the present time it appears to be useful to talk about grammatical cues in this unified way, as Givón does. Grammatical morphology will be discussed in detail in the following chapter. It is perhaps unlikely that all of the varieties of grammatical meanings that are commonly associated with grammatical morphology in the languages of the world are also capable of expression by means of word order. For example, I am unaware of any language in which word order is used to signal person and number categories. However, many grammatical meanings which are morphologically expressed in one language are syntactically expressed in other languages. A commonly cited example is the marking of the grammatical relations of nominal expressions to the verbs with which they are associated (e.g., subject, object, indirect object). A language may use case-marking, adpositions, verbal agreement, word order, or some combination of two or more such devices to signal these relationships. Grammatical agreement in general, like linear adjacency in general, must help to group nouns with their satellites or verbs with their arguments, thus identifying groupings that are relevant to conceptual processes, such as conceptual mergers. Other grammatical meanings are less commonly expressed syntactically. An example is interrogativity. Commonly yes-no questions are marked by interrogative particles or affixes in the languages of the world, but are marked by word order in English and some other related languages. Therefore, it seems reasonable to group syntactic constructions with grammatical morphemes in making reference to a class of *grammatical* cues in addition to *lexical* cues.

Grammatical cues appear to trigger at least six different families of processes. The ones that Givón (1990) discusses in connection with the above quote have to do with (1) attentional activation and (2) search in memory storage, which he relates specifically to the topic of referential coherence. It seems clear that grammatical cues are also involved in (3) the construction of complex conceptual representations, such as a scene in a mental model, from smaller components. It is to this function which Morrow (1986) pointed in saying that grammatical morphemes "organize objects and actions into situations". Next there are those varieties of grammatical morphology that relate to (4) the reality status or epistemic status of the situations modeled by the mental model. In this category we might group negativity, modality and evidentiality and interrogativity. (5) There are markers of interclausal relations, which indicate relationships between mental models such as cause-effect, reason-result, modifier-head, etc. Finally there are (6) aspects of social relationships that may be grammaticalized (honourific systems, deontic modality and imperativity).

1.3.4.1. Syntactic cues

The topic of the linear ordering (including hierarchical grouping) follows

naturally from the topics of speech perception and lexical access. In a sentence with a single adjective and two referentially distinct nouns, for example, the concept triggered by the adjective may be merged with the concept triggered by either of the nouns (for example, in a sentence referring to a man and a tree, and using the adjective *tall*). Which? There may be morphological cues signaling which concepts are to merged, such as gender agreement (where *tall* might be marked as masculine or animate, for example, agreeing with *man*, but not with *tree*) or case concord (e.g., when the man is chopping the tree, and *tall* is marked with the case of subject rather than the case of the direct object). But the combination of linear adjacency and right-left ordering may also be important cues (as when adjectives proceed nouns and we find *tall man tree* and not *man tall tree*). In fact, in natural languages there frequently is a close relationship between linear groupings of words and conceptual operations, such that the linear groupings can be cues to the conceptual operations, as when the linear relation in *tall man*, signals conceptual combination required to produce a concept of a "tall man".

However, even in languages that depend heavily on word order as a grammatical device, there can be instances where the grouping of component concepts into complex concepts is not completely determined by linear ordering. Such *parsing ambiguities* have played a major role in research on syntactic parsing. The assumption has been that that the strategies used in coping with parsing ambiguities should yield insights into the nature of the parsing processes and their relationships to other processes. A common example (from Rayner, Carlson & Frazier, 1983) is the ambiguity seen in the segment *The performer sent the flowers...* in which sent may be the main verb and *the performer* may be the subject. Alternatively, *sent* may be the verb of a reduced relative clause in the passive voice, in which case *The performer sent the flowers...* is synonymous with *The performer who was sent the flowers....* Rayner et al. (1983) presented evidence that during reading, the former parsing choice is consistently made first. That is, an initial sequence of NP-V is taken to be a subject and verb. Rayner et al. (1983) argued that syntactic phrase structure considerations alone control the initial parsing choice and other factors, such as semantic plausibility, only play a role later, if it becomes necessary to reparse the initial NP-V differently in order to create a plausible interpretation, or to exhaustively parse the sentence.

What causes the parser to decide it has made a mistake? Rayner et al. proposed that there is a separate process going on in parallel (and in isolation from the syntactic parsing process) that involves assigning thematic roles to noun phrases. Those thematic roles may turn out to be in conflict with the thematic role assignment required by the syntactic parse, thus forcing a reanalysis.

A considerable amount of effort has gone into finding evidence that purely structural parsing preferences of the sort claimed by Rayner, et al. could be overridden by other factors such as context, the typicality of a noun as an agent or patient, the frequency of a word's occurrence in a particular syntactic configuration and so on (Tanenhaus & Trueswell, 1995). There is some evidence that such factors can override the structural parsing principles, so that a wrong parse is avoided. Altmann (1997) claims that

> The general consensus now is that in fact a whole variety of factors influence the decisions that have to be made when an ambiguity is found. The fit with context is one of these factors, but so it seems is the frequency of occurrence of the different structures associated with the ambiguous words.

However, Frazier and Clifton (1996) have claimed that the manipulations of such variables, although perhaps able to eliminate the effects of purely phrase-structural

factors, have not generally been able to reverse them. If this is so, then certain structural factors (adjacency and linear order) would appear to influence parsing very powerfully, at least in languages like English (Frazier and Clifton, 1996 cite similar findings for Italian, Spanish, German, Dutch and Japanese).

In any case, the processor appears to make parsing choices (decisions regarding how lexical and grammatical cues are to interact with respect to further processing) rapidly and automatically under the influence of various factors. In most cases the initial rapid parsing choices work. When they do not, the processor is usually able to recover from the resulting glitches, though not without extra processing cost.

Not a lot is known yet about the nature of L2 parsing. Juffs (1998a) investigated the L2 English processing of parsing ambiguities of the type discussed above in which an NP-V sequence is ambiguous between a subject-verb parse and a parse in which the V belongs to a passive relative clause modifying the NP. One particularly interesting feature of Juffs' study was the inclusion of postverbal prepositional phrases, as in *The birds killed in the garden....* Juffs reasoned (citing MacDonald, 1994) that for native English readers the postverbal prepositional phrase would force a passive reading of *killed*, since such prepositional phrases (and adverbials generally) cannot intervene between a verb and direct object in English. In other words, in English, a postverbal adverbial is a highly reliable cue to the intransitivity of the verb (passive verbs being only a special case of intransitive verbs). It would be of much interest if it could be determined whether L2 English users eventually acquire parsing strategies such as that involved in using a postverbal adverbial as a cue to intransitivity, and if so, how readily they acquire them. By my understanding, Juffs' findings if anything suggest that the postverbal adverbials did not help these relatively advanced L2 English users.[6]

If this is the case, then at least some parsing cues would appear to be rather late-acquired in L2 English, if in fact they are acquired at all.

1.3.4.2. Morphological cues

We have seen that the mental language processor appears to react rapidly and blindly to the presence of stretches of sound compatible with specific words, identifying as many words as it can and activating their meanings (without conscious awareness on the part of the listener), regardless of any particular word's relevance to the context. There is a further question regarding the identification of portions smaller than words, in particular, roots, stems and affixes. The following chapter deals in detail with grammatical morphology. Here it is worth noting that the general direction of research on the

---

[6] Juffs' Figure 6 would appear to represent the sentences where the verb itself is ambiguous (could be simple past or past participle), and where there is a postverbal adverbial (prepositional phrase). Apart from the postverbal adverbial, native readers might garden-path at the verb, in which case the disambiguating region would be around the words *were playing* in the sentence *The boys criticized during the morning were playing in the park.* The fact that native readers show no decrease in reading speed in that region suggests that they were not garden-pathed. However, all groups of nonnative readers do slow down in the disambiguating region. It appears to me that this would be the crucial case for determining whether the nonnative readers make use of a strategy that treats postverbal adverbials as online cues to the intransitivity of immediately preceding verbs. And it appears that they do not. Juffs argues from other evidence that his L1 Romance participants have acquired sensitivity to the postverbal adverbial as a parsing cue. However, his argument crucially includes cases (his Figure 5) where all nonnative readers perform in a clearly nonnative manner. This makes that particular finding questionable, in my opinion, as a crucial case for arguing for the presence of a nativelike parsing strategy. Rather Figure 6 appears to me to straightforwardly indicate that no group of nonnative readers was helped to avoid the garden path by the presence of the postverbal adverbial, while native readers were helped.

recognition of words in the sound stream, as discussed above, would seem to logically entail sub-word processing in some sense. Upon encountering an utterance containing *beating*, the word *beat*, that is, the base-form of the set word-forms to which *beating* belongs, will be activated. Unlike the homophonous *beet*, the conceptual form of *beat*, will not subsequently be inhibited. If the processor, in addition to identifying and activating *beat*, also identifies the word *beating*, then further processing will be triggered on the basis of the fact that the ultimate word is not *beat*, but *beating*. In other words, there would be two processing steps triggered by *beating*, one caused by the stem portion and the other caused by the deviation of the word from the stem. This would appear to make the question of whether *beating* is parsed into *beat* plus *ing* somewhat difficult to pose. In many cases, however, the realities will be more complicated than this and discussion is postponed to Chapter 2.

In Chapter 2 we will also review research on the acquisition of grammatical morphology in second language acquisition, and in Chapters 3 and 4 we will explore in depth the acquisition of L2 inflectional morphosyntax. To anticipate a bit, we have already seen hints of what we might expect to find. In connection with speech perception, lexical access and some parsing strategies, what little evidence we have found appears to encourage us to expect L2 acquisition to occur, to the extent that it does occur, through gradual strengthening, rather than through discrete-step hypothesis formation, hypothesis revision, or the "triggering" of changes in an internal grammar by limited evidence in the input. It is reasonable to predict similar findings in the domain of inflectional morphology, given that the cue-to-process relationships are at least as complex as those in these other domains.

1.4. COMPREHENSION IN SUMMARY

Returning again to the tourist group with whom we began, we now have filled in more details of what they have that I lack which makes their experience of hearing their language so profoundly different from my experience of hearing it. Basically, what they have is a system of internal reactions to elements of phonetic form and further reactions to the results of earlier reactions. The listeners receive acoustic cues. Their language processors immediately act on those cues, creating the impression of specific familiar speech sounds. With little time lag, their processors react to numerous stretches of sound, recognizing them as contact representations of lexical entries, and immediately activating the meanings of those lexical entries. Contextually inappropriate lexical choices are then eliminated, as the selected lexical meanings are integrated into evolving mental models, with the assistance of grammatical morphemes, prosodic cues and word-order cues. As I listen to their language, I lack both sensitivity to the cues and the specific processing reactions that those cues trigger for the native listeners.

This is a rather different picture from the traditional picture that takes for granted the objective existence of sounds (vowels, consonants, tones, etc.), morphemes, words, phrases and so on as the building blocks of language and the substance of language learning. The fact that the Russian examples discussed below are presented in orthographic form may encourage us to lose the point that learners are working from an acoustic stream, and mentally creating whatever sense they have of discrete linguistic units. The basis for successful creation of this sense of linguistic units is in part language specific, that is, in and of itself includes an important part of what must be learned, along with deeper aspects of the cue cascade. In Chapter 3 we will want to see whether learners have acquired sensitivity to inflectional cues, and whether they have acquired the nativelike associations of those cues with components of comprehension processes. Do inflectional cues derivable from the sound-stream cause any reactions from the L2 processors? If so,

does the variety and strength of the reactions they cause change over time; that is, is there a pattern according to which learners acquire the ability to react to inflectional cues in L2 Russian?

1.5. LANGUAGE PROCESSING AND LANGUAGE LEARNING

The above discussion was aimed at giving some substance to the notion of the *what* of second language acquisition. That which is acquired includes sensitivity to cues and processes triggered by the cues, leading to the formation and evolution of mental models. (Needless to say, the discussion has been extremely sketchy, omitting whole areas of language processing and skirting controversies.) That does not address the *how* of second language acquisition, but it does simplify the job of answering this additional question, in an even sketchier manner! In the previous paragraph, I referred to the *association* of cues with aspects of comprehension processes. In this section we consider the possibility that learning consists in the formation of such associations. In the discussion of mental models, I argued that the ability to construct them exists independently of the use of linguistic cues to trigger and constrain their construction. Learning is possible as long as there are parallel information sources (Klein, 1986), with one source providing the basis for mental model construction and the other providing the linguistic cues that need to be learned. In early language learning, the mental models running in parallel to the speech flow might be based on direct perception of what is being talked about (here-and-now language). Later, the learner may be able to construct models that are less dependent on the speech flow, based on a combination of lexical information and previous experience, as when a child L1 learner hears his or her caregiver relate events in which s/he personally participated. Later still, a combination of lexical information and inferential processes may provide adequate mental models of text, even though the learner is not yet able to make use of all of the cues provided within the text.

As long as a conceptual stream is created in parallel with the sound stream, there is a chance of associating events in the two streams with one another. Suppose it happens to be the case that often when a dog token is added to a mental model, the sound segment corresponding to the word *dog* is also present in working memory.[7] This would favour the association of that small, recurring stretch of sound pattern with the concept of 'dog' (rooted in the learner's experience of dogs). Once the association is strong enough, the sound pattern, which was originally *independently* found to co-occur with the concept, will be able to *trigger* the concept, that is, to cause a dog token to be added to the mental model.

Such associative learning should take place most easily in the case of concrete lexical meanings. Concepts of dogs, of whiteness, of cats, of blackness and of the activity of chasing can presumably be maintained with effort in working memory over some period of time, while the sound patterns needing to be associated with them can also presumably be held in an articulatory or (perhaps better) auditory loop (Gupta & MacWhinney, 1995). This potential time-stability of both the concept and the acoustic substance could allow an association to form relatively readily. Unlike such lexical meanings, grammatical "meanings" relate to events involved in the construction of mental models (and other properties of mental models referred to briefly above). The

[7] The notion of working memory assumed here is the one expressed by Kosslyn (1996): "Working memory includes the information being held in the various short-term memory structures plus the information that is activated in the various long-term memory structures. In addition it includes the 'control processes' (including the property lookup subsystems) that activate information in long-term memory and maintain information in short-term memory." (p. 324)

event of merging the dog concept with the whiteness concept, or the cat concept with the blackness concept, might not have the much potential for extended temporal duration in working memory, making associative learning of related grammatical cues more difficult than associative learning of lexical cues. If so, then in order for strong associations to take hold between formal cues (e.g., word order, agreement marking) and such conceptual merger events, the learner might need to experience their co-occurrence many more times than might be necessary for learning lexical meanings. Or consider the "meaning" associated with the English perfect aspect, as in *Annie has met Fred.* On the assumption that the auxiliary *has* is triggering or constraining comprehension processes, let us hypothesize that these processes have to do with the facts that 1) the action is not in the stream of events of an evolving narrative model and 2) there is a need for a heightened degree of persistence of activation in memory to be assigned to the representation of Annie's meeting Fred. That is, perfect aspect is used when the content of the clause is especially relevant to further inferences, increasing the readiness with which such inferences would arise. Associative learning in this case would require the repeated co-occurrence of contexts in which the perfect marking occurs in parallel with an action that is not in the chain of events constituting the narrative model, and in which the clause so marked is highly relevant (somewhat in the sense of Sperber and Wilson, 1986) to further inferential processes. Once formed, this association could subsequently serve to prevent the addition of a perfect-marked event to a narrative model, and cause a heightened level of activation (or at least more enduring activation) of the information derived from the clause. We should not be surprised to find that this association would not be detected without there being a large amount of language processing ability already in place, after which point it still might require a considerable amount of experience the association to become a strong one.

In the case just hypothesized, some properties of mental models (including their relevance to inferential processes) would need to be categorizable, as would the sound pattern associated with those properties. This raises the question of which aspects of mental models and the mental model construction process are categorizable, and hence, can become lexicalized or grammaticalized in languages. Languages themselves provide us with the primary evidence in this regard. This will be discussed further in the following chapter. Briefly, any event of mental model construction that is grammaticalized in a wide variety of languages must be assumed to reflect a type of step(s) in mental model construction that is (are) both categorizable and reasonably salient. Take the merger of actions with the patients of those actions (for example, merging the concept of *washing* with the concept of *dishes*, or with the concept of *hands*, or the concept of a *shirt*, or the concept of a *window* or of a *car*). The concepts to be merged are typically flagged grammatically by case marking, word order, verbal agreement, etc. Besides that, verbs subcategorized for optional direct objects are often marked by verbal morphology indicating the *intransitivization* of grammatically transitive verbs or the *transitivization* of intransitive verbs. In fact such markers of valence changes are very widespread cross-linguistically (Bybee, 1985). Thus it appears that the merger of action and patient (and related mergers) is a major, salient, readily categorizable step in mental model construction, and languages will commonly grammaticalize markers of its occurrence. To personify, we might say that human language processors like to be able to know whether or not to anticipate merging a patient with the action indicated by a verb.

At the outset of this chapter we contrasted two possible (families of) views of language acquisition. On the more common view, what is acquired in SLA is a grammar, that is, a system which provides the language user with information regarding the form of sentences. The other view takes the elements of linguistic form to be primarily processing triggers. It was suggested that even if these orientations are not empirically incompatible,

they nevertheless would tend to influence the approach to specific questions, and influence the search for relevant data. I would like to illustrate this difference in orientation briefly by turning to an oft studied issue in the acquisition of English as a second language: the placement of adverbs in relation to verbs.

Nonnative English utterances of the form *Mary watches often television* have especially attracted attention because they are believed to reflect a parametric difference between English and other languages (especially French) and are thus believed to provide a testing ground for hypotheses which frame second language acquisition in terms of parameter resetting within the Principles and Parameters orientation to syntax (White 1989, 1991a, 1991b, 1996, Trahey and White, 1993). The putative parameter involved has changed over the years. An early version (White, 1989) involved the parameter of Adjacency of Case Assignment: English, but not French, requires that the verb and the direct object to which it assigns abstract case be adjacent. The current version involves the "strength" of agreement features (strong AGR vs. weak AGR) with the concomitant presence or absence of overt verb raising. Thus *Mary watches often television* has the verb *watches* in the raised (fronted) position, even though in native English, such verb movement is covert (not visible) due to the fact that AGR is weak in English. Eubank and Grace (1998) allow that the setting of the relevant parameter might be learned on the basis of observed word orders (although they would prefer to find that the acquisition of English inflection will automatically provide the information on the strength of AGR, resulting in the correct parameter setting).

In any case, the assumption seems to be that successful learners must acquire knowledge of English such that they end up with target-like word order (*Mary often watches television*). Regardless of the degree of theoretical sophistication, the goal of learning in such an account would appear to be for the learner to reach a state in which the verb and adverb are always placed in a nativelike linear ordering relationship, without apparent concern for the effects of the alternative orderings on comprehension processes.[8]

Thus, White and collaborators have studied various strategies (that ultimately amount to telling learners, "In English, we don't say *Mary watches often television;* we must say *Mary often watches television*") in hopes that this information will cause a modification of the internal grammar such that the offending word order will no longer appear in the learner's spoken production.

Taking the alternative perspective, that is, viewing the details of linguistic form as processing cues, we ask what different comprehension process(es) are triggered by the alternative word orders. It would appear that in normal native English speech, the immediate postverbal placement of an adverb provides an extremely reliable cue to verb valency (see section 1.3.4.1. above), which as we have already noted from a cross-linguistic perspective, appears to be an important bit of information to human language processors. (That amounts to saying that the conceptual step of merging an action with a patient makes a relatively big splash on the conceptual plane.) Therefore, on hearing *Mary watches often...* the native English parser is in a position to conclude that there will be no

---

[8] It is important to keep in mind that that the issue is one of the native comprehension *processes*, and not the *ultimate understanding* of such sentences. In terms of the ultimate understanding, the word order may appear to be irrelevant, while in terms of the mechanisms that lead to that understanding, nonnative word order should cause a particular processing glitch. Traditional (global) measures of comprehension, such as comprehension questions, do not provide information regarding the extent to which the specific mechanisms of comprehension are nativelike or nonnativelike. Yet nonnative like features of spoken production go wrong at the level of those specific mechanisms, and typically not at the level of global comprehension success.

direct object in the sentence and it apparently draws this conclusion. The sentence might continue ...*but participates rarely*. If for native speakers, the order *Mary watches often* leads the parser to conclude that there will be no direct object, then a sentence such as *Mary watches often television*, containing a direct object, will provide the mental language processor with clashing comprehension cues. What is lacking in the nonnative English user who produces *Mary watches often television* need not be abstract knowledge related to the proper arrangement of words and phrases in English, but rather a particular use of a word order cue to constrain the parsing process in a particular way. Were the nonnative speaker to acquire the relevant parsing strategy, we might expect sentences of the offending variety to gradually disappear from his or her spoken productions, since they would strongly clash with his or her own parsing strategies.[9]

It is not at all obvious that relatively brief pedagogical intervention would effect the development of a parsing strategy. We might doubt that simply teaching learners what one can and cannot say in English would readily lead to the development of such a parsing strategy. Rather, the parser might require an enormous volume of experience before it would learn that a postverbal adverb is a highly reliable cue to the absence of a subsequent direct object. Thus, viewing the elements of grammar as processing cues not only points us to an alternative account of the persistence of the nonnative production patterns, but suggests that in such cases, traditional pedagogical interventions will not achieve rapid, easy success (as might be expected in particular, if the relevant learning involved the triggering of a parameter setting by positive and/or negative evidence, or the movement from a stage i to stage $i + 1$ caused by input for which the learner was in a state of readiness).

1.6. SPEECH PRODUCTION

Viewing SLA as the acquisition of a cue system (consisting of the cues themselves and the processes triggered by those cues) could not provide a total account of SLA, since a total account must also include the development of a speech production system (and yet other systems for reading comprehension and written production). The challenge for the spoken production system is to produce cues in compliance with the needs of the interlocutor's comprehension system. It would seem that the bigger challenge in language learning would be the acquisition of the system of comprehension cues. This aspect of learning depends, it was suggested above, on the association of elements of linguistic form with components of conceptual representations, with steps in the construction of conceptual representations, and with certain other properties of conceptual representations (such as their epistemic status). Once some of these cue-process associations have developed, there will be a learner-internal basis for the development of a spoken production system. Utterances produced will succeed or fail to meet the needs of this comprehension system to varying degrees. It is conceivable that a fairly general learning system, behaving in the broad spirit of a connectionist network (Rumelhart & McClelland, 1986), could come to produce utterances with the necessary

---

[9] It should be clear that this approach to "ungrammaticality", i.e., in terms of clashing comprehension cues, provides an alternative approach to the logical problem of language acquisition. Processing cues and their associations with specific processes can be acquired on the basis of positive evidence alone. Once cues have become strong, certain combinations will be strongly incompatible and certain reliable parsing strategies will fail. It is an interesting exercise to apply this concept of ungrammaticality (as consisting in cue clashes and parsing failures) to various classic varieties of ungrammaticality. See Juffs & Harrington, 1995 for evidence that L2 ungrammaticality judgments reflect parsing problems.

qualities. The stimulus triggering production would be a conceptual stream of the same nature as the conceptual streams that are the end-product of comprehension. That is, comprehension and production would share a common message level, or level of mental models. Production patterns (or subpatterns) which smoothly click through the learner's own comprehension system could be strengthened, while those that clash with it, for example, by providing conflicting cues, could be weakened. (The idea of the comprehension system training the production system in this way was suggested by Stemberger, 1998.)

On this account, in terms of normal language mechanisms, the production system could not be more nativelike than the comprehension system. Given that learning must take place in time, we can imagine the production system at a given point not yet reflecting the current state of the comprehension system. That is, we might in principle expect that in comparison with the production system, the comprehension system would always be either more nativelike than the production system, or at least equally nativelike. To the extent that L2 utterances reflect features of nativelike form in advance of the corresponding developments in the comprehension system, this would by definition involve cognitive strategies other than normal linguistic speech production processes.

The relationship of "errors" in the production system to cue processing in the comprehension system must be complex. For example, if an L2 English learner produces *I have lived here since three years*, a traditional account might hold that the learner has not yet learned to use the particle *for* as required by English grammar to produce the nativelike *I have lived here for three years*. However, what is missing from the comprehension system is not necessarily related to the "correct" target *I have lived here for three years*. Rather, the fact that the error goes unnoticed by the learner-speaker suggests that the particle *since* is not being processed in a nativelike way in the learner's own listening comprehension system. The particle *since* in native English requires a particular type of temporal complement. That temporal complement must be conceivable of as a starting point in time, not a span of time, (as in *I have lived here since 1997*). The native English language processor, on encountering since would begin to anticipate such a complement. That is, the native processor, on encountering *I have lived here since three years...* might expect a continuation such as *...ago*. This hypothetical L2 user's processor appears to lack such an expectation. In the event that the L2 user's processing of *since* were to come to conform to the nativelike pattern of expectations, productions such as *I have lived here since three years* would begin to clash with the expectations of the learner's own comprehension system, leading to an abandonment of such productions in favour of others which do not clash with the production system. At that point, the competing particle *for* could begin to win out over *since* in the production of such utterances.

In the end, the production system as "trained" by the comprehension system need not have any resemblance to the comprehension system in its overall form and organization, other than at the level of mental models. Bock and Levelt state that there is "reasonable agreement on the broad outline" of steps in the speech production process (Bock & Levelt, 1994, p. 945). Following Garrett (1982, 1988) they propose four processing subcomponents which convert messages into spoken output. Following its initial formulation, the message is next processed at the *functional level*. At this level specific words are selected (but not phonologically encoded) and assigned their grammatical functions such as subject and object (but without being sequentially arranged and without grammatical morphology). At the subsequent *positional level*, words are arranged sequentially and inflectional forms chosen. A final level of *phonological encoding* gives utterances their actual spoken form. The comprehension system, as we have been picturing it, does not precisely mirror this order of processes. Rather, it from

acoustic form to word identification to positional relationships. In any case, the time constraints of speech might favour a fast and dumb production system, the learning of which is controlled by reactions of the comprehension system, but which then takes on a life of its own. As long as the production system learns to convert messages into forms that work smoothly for the comprehension system, that is enough.

1.7. THE REDUNDANCY OF THE *INTERNAL GRAMMAR* NOTION

From the perspective of the processes involved in comprehension and production then, we might wonder whether there is such a thing as an internal competence grammar at all. The child confronted with the language in its environment is initially in the position that I was in when listening to an unknown language at the airport. There is a wall of sound within which acoustic cues must come to be detected. A cascading cue-process system must develop in which acoustic cues indirectly trigger and constrain the formation and evolution of mental models. The child must also develop a production system that provides the very cues that the comprehension system has learned to process. If the events of mental model formation that trigger spoken production are the same as those that occur during comprehension, then those conceptual events can themselves become the cues that trigger the cascade of processes involved in production. These comprehension and production mechanisms are what are generally thought of as performance systems. What purpose would a separate internal competence apparatus serve?

Traditionally, descriptive grammars were the result of applying nonlinguistic cognitive capacities (no doubt with some dependency on language processing mechanisms) to the examination and analysis of the output of language production systems. There is a natural distinction between the observable properties of the products of some behaviour (such as a beehive, a spider web, or spoken utterances and discourses) and the internal "instructions" underlying the behaviour itself. Presumably the system within each honey bee which leads to the steps necessary to construct a hive would not best be modeled by a description of the geometrical properties of the hive. It is somewhat puzzling how, in the case of human language, descriptions of the product continue to be treated as models of the internal apparatus underlying the process.

1.8. CONCLUSION

In this chapter, I have attempted to make the point that viewing SLA as grammar learning, aimed at acquiring static knowledge of the form of grammatical sentences (knowledge which also excludes ungrammatical sentences) is only one option available to SLA researchers.

Another option is to assume that what is acquired are the comprehension and production processes which make it possible for language users to convert a complex sound stream into understanding, or a convert a stream of understanding into spoken articulation. In both comprehension and production, the processes involve cascades of computations of enormous complexity taking place at dazzling speeds and with seeming transparency at the level of the language user's subjective experience. On this view, the form of language is fundamentally related to comprehension, with formal elements functioning as cues which trigger and constrain aspects of the comprehension process.

What is to be learned then, is not a set of ready made units such as phonemes, morphemes, words, or phrase patterns, sitting there like fruit ready to be plucked from the input and added to the internal knowledge system. What evidence is available, as we noted, may indicate that the development of consistent detection of acoustic cues, the

development of a mental lexicon with strong, immediate connections between auditory contact representations and conceptual representations and the development of strong parsing expectations are dependent, not on categorial, discrete learning experiences, but rather involve mental associations which in many cases take hold with difficulty and are then gradually strengthened over a long period of time.

We have also hinted that the same may be true of the second language acquisition of grammatical morphology in general and inflectional morphosyntax in particular. Since inflectional morphosyntax is the domain of the experimental research reported and discussed in Chapters 3 and 4, it will be helpful to reflect in more detail on the complexity of the form and functions of morphologically expressed grammatical categories. It is to this sub-system of grammatical cues that we now turn.

2. Inflectional Cues

The speech stream thus presents the listener with cues that trigger and constrain components of the comprehension process, understood as a process of mental model construction and modification. Language learning is a matter of associating cues with corresponding processes. Inflectional morphology has a particular place within the larger cue system. In this chapter, we seek a fuller understanding of this particular part of the overall cue system. This provides the context for the examination of the development of inflectional processing ability in L2 Russian, the topic of the following two chapters. Hopefully, this current chapter will also help us to appreciate why the development of inflectional processing mechanisms may not occur rapidly. After some preliminary remarks, the chapter is divided between a general consideration of the nature of inflectional morphology, some background on Russian inflection, and a review of major strands of research on the second language acquisition of inflectional morphology.

2.1. PRELIMINARIES

We begin with a general clarification of the domain of inflectional morphology, depending mainly on intuitive examples, but also clarifying key terminology, and then present some basic evidence that encourages us to think of inflectional form as a variety of processing trigger.

2.1.1. Definitions

In a discussion of inflectional morphology, there are various possible sources of terminological confusion, especially related to the term *word*. In traditional terms, someone might say, "*Bake, bakes, baking and baked* are different forms of the same word." What then is the word that they are forms of? Paradigms such as this are based around the intuition that something essential unites them. In processing terms, what unites them is that they all trigger the addition (or continued activation), in one way or another, of the same basic atomic component of mental models. They are united as instances of the same lexical cue, on the basis of their commonality in terms of comprehension processes they trigger. I will follow Matthews (1991) in using the term *lexeme* to indicate what it is that unites the forms of a paradigm such as this. Thus we can say, "*Bake, bakes, baking and baked* are different forms of the same lexeme." (Note that this is not the usage of the term *lexeme* found in some of Levelt's writings, e.g., Levelt, Roelofs & Meyer, 1999, where, following Kempen and Hoenkamp, 1987, the term is used in reference to a word's phonological properties as distinct from the semantic and grammatical properties which Levelt refers to as the *lemma* of the word.)

The term *word* is also used at times for the individual members of a paradigm: "The word *bakes* is a third person singular present habitual verb". For this sense of the term *word*, I will again follow Matthews (1991) in using the term *word-form*. We now can say, "The word-forms *bake, bakes, baking and baked* are forms of the same lexeme".

Following another fairly natural turn of phrase, someone might also say, "*Baked* is a form of the verb *bake*" (or "*to bake*"). This could be taken as equating the lexeme with one of the word-forms. Alternatively, it could be understood that one of the specific word-forms is most naturally taken as "the name of the word", or more precisely the name of the lexeme, i.e., the basic form of the word. This form of a given lexeme, which I will henceforth call the base-form, has various properties. Günther (1988) lists the

following: (1) It is the form that is given when the word is supplied by a native speaker independently of any syntactic context (as in a word list). (2) It is the form most commonly substituted for other forms in agrammatic aphasic speech production (citing the case of German). (3) It is the form acquired first by children. There is also experimental evidence, at least in the case of nouns, for the privileged place of the base-form in lexical access (Günther, 1988; Lukatela, et al., 1978; Lukatela, Gligorijević, Kostić & Turvey,1980; Feldman & Fowler, 1987).

There is potential ambiguity, once again, in the use of the term *base-form*, in that this term is also used by some authors in reference to the stem to which inflectional affixes are added. In English, these two senses of *base-form* happen to coincide (that is, the bare-stem form is also the base-form in the sense just discussed). In Russian, however, this is not always the case. For first declension nouns, the base-form is the unsuffixed form, as with *mal'chik*, 'boy'. In such a case, adding the suffix *–a* to the stem produces a non-base form, the genitive singular, *mal'chika*, 'of the boy'. However with second declension nouns, the base-form, that is, the nominative plural, already has a suffix, *-a*, as in *kniga*, 'book'. The corresponding stem with no suffix is also a word form in this case, but rather than being the nominative (and hence the base-form), the bare stem form is the genitive plural, as in *knig*, 'of the books'. The base-form is also the *unmarked* form, in the sense to be discussed.[10]

These definitions of *lexeme*, *word-form* and *base-form* will prove helpful in avoiding confusion in many contexts. In other contexts where nothing is at stake, the vaguer term *word* will continue to be used.

We will briefly postpone the characterizations of inflectional cues as such, until we have discussed evidence that inflectional cues do in fact trigger extra processing over and above that triggered by the lexemes to which they are bound.

2.1.2. Inflection triggers processing.

In this section, we will see some basic evidence that in normal language processing, grammatical morphology in general, and inflectional morphology in particular, play a vigorous, lively role in languages. This understanding ought to affect the way grammatical morphology is viewed in connection with SLA. For example, if native speakers of English say, *She likes me*, instead of **She like me*, merely because that reflects a formal requirement of English for a redundant and rather useless bit of formal substance, then the L2 English learner's problem boils down to getting the right forms in the right places. On the other hand, if the *–s* of *likes* (or alternatively, the fact that the form is *likes* and not *like, liked,* or *liking*) triggers a flurry of processing events, that puts the L2 learner's situation in a different light.

Within the broader category of grammatical morphemes, inflectional cues differ from function words in that they are tightly fused with lexical cues into fundamental processing units—phonological words (yet another sense of the word *word*). This can be seen as a useful tactic. If the lexical cue provides a building block for mental model construction and the grammatical cues tell the processor what to do with that building block, then providing both in this unambiguously fused manner makes some sense. But it means that an inflected word must trigger two kinds of processes almost simultaneously: the addition of a conceptual atom and the deployment of that atom within a larger

[10] The question of the base-form for verbs is more complicated, involving, I will suggest, the notion of local markedness (Tiersma, 1982).

conceptual molecule (related to the clause or sentence containing it) or macromolecule (related to the discourse containing it).

To be concrete, consider the Russian word *sobaku*, the accusative form of the lexeme for 'dog'. The lexical cue (the presence of a word-form belonging to a particular lexeme) instructs the language processor to reactivate a dog token in the discourse model, or if none is available, to add one. The dog token is a conceptual atom. The accusative form (the ending *–u*, or alternatively, the fact that the word-form is *sobaku* and not the base-form *sobaka*) further instructs the processor that the dog is to be merged with the action expressed by a verb, thus leading to the formation of a conceptual molecule specific to the particular action, the form that that action takes when dogs undergo. The liveliness of this inflectional marking as a processing trigger is made apparent by contrasting normal listeners with listeners for whom this inflectional cue has ceased to function. The Russian Neurolinguist A.R. Luria (1946, cited by Alpatov, 1997) presented the following sentence to aphasic patients:

(1) Sobak-u oblaja-l-a loshad'
Dog-accusative bark.at-past-feminine horse:direct.case
A horse barked at a dog.

The noun *loshad'*, 'horse', belonging to the third declension, does not distinguish nominative case and accusative case (hence the designation *direct case*). The verb *oblajala*, 'barked.at:past:feminine' indicates by its gender agreement that the subject is feminine, but both *sobaku*, 'dog:accusative' and *loshad'*, 'horse:direct.case' are feminine. The verb *oblajala* has a selectional restriction requiring that its subject be a dog (as suggested by the English translation *barked*). On the other hand, it requires two arguments: an actor and an undergoer. The word *sobaku*, 'dog:accusative', by virtue of its case marking, must refer to the undergoer, while the other argument, *loshad'*, 'horse:direct.case', in and of itself, is compatible with either an actor or undergoer role (depending in part on the particular verb). The lexical cues would thus lead us to expect a mental model in which a dog barks at a horse. On the other hand, the grammatical cues require a mental model with a horse barking at a dog. This sentence therefore elicits laughter from audiences of healthy native Russian listeners. Some of Luria's patients understood this sentence as unproblematically meaning that a dog was barking at a horse. For such patients, the case-marking of *sobaku*, 'dog:accusative', was *inert*. Their ability to understand the sentence as a whole suggests that adequate phonetic processing occurred. That is, there is no reason to believe they did not hear the case ending. However, for them it was as though the case-marking were not there. This inflectional cue did not trigger any further processing beyond the processing triggered by the lexical cue.

Experimental evidence that inflectional form triggers mental processing above and beyond that triggered by the lexical content of the inflected word is provided by Thomson and Zawayedeh (1996). Participants in their experiment performed lexical decisions to visually presented words while listening to auditorily presented sentences. The lexical decision targets were semantically unrelated to the sentences. The sentences were either past progressive in form, such as *Just before dawn the soldiers were nervously guarding the entrance to the palace*, or they contained a sequence of three narrative events, as in *The doctor delivered the baby, paddled its behind and listened to it holler*. The lexical decision target appeared on a computer screen at a point in time near the end of the sentence. The targets were either non-words, uninflected verbs, or verbs inflected with either *–ed* or *–s*. The key finding was that uninflected verbs were insensitive to the type of sentence being processed auditorily. That is, the verb *kick* would be responded to as

quickly against the background of a three-event sentence as it would against the background of a past progressive sentence. In the case of inflected verbs (e.g., *kicked*), there was a significant delay in responses in the context of the three-event sentences in comparison with the past progressive sentences. These results were consistent with a view according to which the uninflected verb forms were responded to independently of the sentences being processed, the uninflected form simply being treated as the "name" of a word (Günther, 1988). By contrast, inflected verbs could not be treated simply as words unrelated to context. Rather, it was argued, the inflection forced a brief effort to integrate the meaning of the verb into the mental model. In the context of the experimental task, there was a sense in which the base-forms remained inert to the discourse processing, while the inflected forms were unpreventably alive and active.

The first line of Lewis Carroll's poem *Jabberwocky* illustrates the power of grammatical morphemes generally to give life to utterances, even in the absence of lexical meanings: *Twas brillig and the slithy toves did gire and gimble in the wabe.* An analogous sentence in Russian is attributed to L. V. Sherba (Gorelov & Sedov, 1997):

(2) Glok-aja kuzdr-a shtek-o budla-nu-l-a bokr-a i kudr'achi-t bokr'-onk-a
English gloss: A glok-**y** kuzdor shtek-**ly** budl-**ed** a bokor **and is** kudhr'ach-**ing the** bokor-**ling.**

(English articles, affixes, etc. are intended to give English readers something a feel for the Russian sentence, although English lacks the Russian case, gender, etc. and Russian lacks the English articles, etc.)

It is reported that Russian grammatical morphemes have such power to organize a mental scene from the nonce stems that some small children will attempt to draw a picture depicting the meaning of a text which, like the above example, contains no genuine lexical items but is organized around grammatical morphemes (Gorelov & Sedov, 1997, p. 47). As the first line of Jabberwocky should impress English readers, so this Russian example should impress Russian readers that grammatical morphemes and other closed-class elements are an extremely powerful force in language processing.

2.2. FUNCTIONS AND FORMS OF INFLECTIONAL CUES

We begin with a general consideration of the nature of inflectional cues, surveying their range of functions and considering their formal nature.

2.2.1. The nature of grammatical meanings

In this section, we are concerned with the range of *functions* of grammatical morphemes observed cross-linguistically. Are those functions consistent with the idea of grammatical morphemes as processing cues that work together with lexical cues in the way discussed above? We also consider a special property of grammatical morphemes: the frequent presence of multiple functions. This multifunctionality presumably raises challenges for acquisition.

2.2.1.1. Crosslinguistic survey

To view inflectional markers, or grammatical morphemes in general, as processing cues is to treat them as two-part entities: a formal cue and an associated process (or set of

processes) triggered or constrained by the cue. We noted above that a cross-linguistic survey of morphological grammatical cues could be thought of as a survey of the types of processes involved in the construction of mental models out of the concepts triggered by lexical cues.

Bybee et al. (1994) sought to catalogue examples of grammatical meanings found in a carefully selected sample of seventy-five languages taken to be representative, areally and genetically, of the world's languages. The authors concerned themselves only with verb-related closed-class elements. Closed-class morphemes were thus included in the database if they had a fixed positional relationship to verbs.[11]

Labeling the grammatical meanings found in the languages of the sample required 358 "meaning labels". That might give the impression of a huge range of grammatical meanings. In fact the frequently recurring meaning labels can generally be grouped into those related to **tense and aspect**, those related to **modality** (including evidentiality), those which **relate verbs to their arguments** in various ways, those related to **illocutionary force** and those which **relate clauses to other clauses** in various ways.

A number of the meaning labels which occurred only once were spatial-locational-directional in nature (e.g., *east*, *center of the lagoon*). A few appear to have manner adverbial meanings (e.g., *clandestine*, *gently*) and others are harder to classify but correspond to small classes of free forms in other languages, such as quantifiers (e.g., *some*, *many*). Most of these less common meanings would not be considered inflectional and perhaps do not even fall into the category of grammatical meanings. Rather, "particles" and "adverbs" of various sorts happened to meet the criterion of belonging to small closed classes and being positionally fixed in relationship to verbs.

Although the GRAMCATS survey involved only verb-related grammatical morphemes, some of these same categories we can expect to find frequently associated with nouns and their satellites. For example, markers of number (singular, plural, etc.) are widespread in association with both verbs and nouns (pointing the conceptual salience of the difference between singular mental model tokens and tokens for collectivities). Anderson (1985) lists the following as common inflectional categories of nouns: number, gender, inflectional class, definiteness and diminution (as "inherent" categories) and case (direct and oblique) and possession (as relational categories). Stump's (1998) list is similar. Inflection of noun satellites (e.g. adjectives, demonstratives) tends to involve agreement with nouns (in either inherent features such as gender or relational features such as case).

2.2.1.2. Functions of case inflection and tense/aspect

It turns out then that the varieties of frequent grammatical meanings (that is, functions of grammatical morphology) boil down to a small number. Do they always appear to relate to components of mental model construction and modification as proposed above? In this regard we might look for at least two types of cues: those related to the building of individual "frames" in the evolving "mental movie" and those related to the changes going on from frame to frame (and connecting the various parts of the movie as a whole). Inflectional morphology which relates verbs to their arguments (e.g., case and verbal agreement) falls into the first category (more on this below). Tense-aspect morphology falls into the second category. These are the primary two functions of inflection that are involved in the categories explored in the experiments reported and discussed in the following two chapters. The basic functions of case marking (and verbal

[11] I would like to thank Joan Bybee for supplying me with the full list of meaning labels and their overall frequencies of occurrence within the GRAMCATS database.

person agreement) are not particularly controversial and will be discussed only briefly. The issues related to tense-aspect marking are relatively more complex and, given the place of tense-aspect marking in the experiments reported and discussed in the following two chapters, it will be worth taking a little more space to spell out certain assumptions which will, it is hoped, help to account for the prolonged time-frame of the acquisition of the formal expression of these categories in L2 Russian.

2.2.1.2.1. Case inflection

For both case morphology and verbal agreement, one obvious function is to relate verbs to their arguments. In looking at isolated sentences, marking of nominative and accusative case, for example, may appear to simply tell the listener who performed the action on whom/what. However, in connected text, the picture is more complicated. Givón (1983) suggests that we "must consider the case-marking of *subjects* and *direct objects* in general ... as part of the grammar of topic continuity" (p. 38). Givón (1984), like many others, makes an important distinction between referential semantic roles (agent, patient, experiencer, instrument, location, etc.) and what he calls pragmatic roles (primary topic, secondary topic), that is roles that have to do with the relationship of the referent to the focal center of the mental model. I will generally speak of the latter as discourse roles rather than pragmatic roles, although it should be understood that even single utterance discourses have a focal center. That is, while some discourse functions may only come into play when multisentence texts are being processed, these discourse roles of noun phrases appear to be always involved in sentence processing. The grammatical functions of subject and to a lesser degree, direct object, involve an intersection between discourse roles and semantic roles. Thus nominative marking may have a primary function of indicating that a participant is the primary focal participant. Only in combination with other cues (verbal valence, voice, etc.) may it indicate the semantic role of the subject (agent, experiencer, etc.). It is conceivable that one of these two functions of case marking would be acquired before the other. In Chapters 3 and 4 the issue of whether the discourse functions of case are acquired before the semantic role-marking functions will be discussed.

2.2.1.2.2. Tense-aspect inflection

From an analytical perspective it is probably most popular to relate tense and aspect to *time*, tense having to do with the "location" of events along the temporal dimension and aspect having to do with other temporal aspects of events , such as their beginnings, ends, whether they occur over a period of time or momentaneously, whether they occur repeatedly over time, etc. (Comrie, 1976, 1985, Dahl, 1985). For someone approaching language as an observer puzzling over the question of the meaning of tense or aspect markers, it often seems natural to conclude that a past tense marker is there to convey to the listener the information that the action referred to by the verb occurred prior to the time of speaking. However, from the standpoint of the role of tense-aspect markers in comprehension processes, one would have to wonder why such temporal relations would come to be obligatorily expressed in every finite clause. Languages do have lexical items with temporal meanings. Languages have temporal orienters such as *one time, the other day, while on my way home from work*, etc. It is common in the languages of the world to employ such devices to situate reports of displaced experience somewhere in time at the beginning of a narrative. For tense and aspect to warrant obligatory expression in a large portion of the world's languages, we should expect them to have a function that is more valuable in sentence-by-sentence processing than merely providing constant reminders that an event or activity took place prior to the time of the speech act.

One arguably more useful function of tense and aspect that has been proposed is the function of indicating when narrated events are supposed to move the narrative model forward to a new point versus when the activities referred to are to be taken as "staging" for other events, or to be taken as general facts about characters and so on.

More generally, a speaker may be narrating events not experienced by a listener, or describing events being observed by the listener, or expressing personal intentions, requesting actions on the part of the listener and so forth. For each of these language functions, the temporal details of pastness, presentness and futurehood seem to follow almost incidentally, hardly meriting ubiquitous obligatory verbalization.

The emphasis on tense-aspect as marking foreground (events that move the narrative forward a step) and background (activities and states that set the stage for foreground events) is especially due to Hopper (1979, 1982; Hopper & Thompson, 1980; in relation to Russian, see Chvany, 1985). Although there is considerable evidence that early uses of aspect marking in both L1 acquisition and L2 acquisition are closely tied to the "aspectual semantics" of particular verbs (see the section below on SLA tense and aspect), once narrative ability is acquired, aspectual distinctions no longer relate in an absolute manner to the nature of the real-world event or activity reported, but rather to the role of that event or activity in the development of the discourse model. Thus *John struck the match* and *John was striking the match* can be used to describe the same state of affairs. It is common to say that the imperfective or progressive version, *John was striking the match* designates a process without its endpoints, while the perfective (simple past) version *John struck the match* includes the endpoints. In fact, the imperfective variant is not conceptually incompatible with the inclusion of the endpoints. However, these two sentences do have different impacts on the mental model that follows their comprehension.

The impact of English imperfective (progressive) form and perfective (simple past) form on evolving mental models has been investigated experimentally by Morrow (1985, 1990). After hearing sentences of the form *John walked through the kitchen into the bedroom*, particularly if such sentences were presented in the context of a developing narrative, listeners indicated that in their model of the situation, John was located in the bedroom following the conclusion of the sentence. On the other hand, after the conclusion of a sentence of the form *John was walking through the kitchen into the bedroom*, the mental model of the situation would have John located in the kitchen. The only difference in the two cases is the verbal aspect, but the effect is for listeners to construct different mental models with the protagonist in different locations. Perfective forms lead to particular modifications of the text model, as seen in Chapter 1, when a Russian reference to opening a car, marked for perfective aspect, led to a mental model in which the car's door was standing open, a state of affairs which subsequently clashed with the further instruction to open the car door.

From the present perspective then, rather than conveying information regarding *when* states of affairs obtain, tense-aspect marking is triggering or constraining modifications to mental models of discourses. Perfective marking, for example, increments the model to a scene further in the development of the model than the scene that would follow a parallel imperfective-marked description of the same event or activity. Each type of aspectual marking can be viewed as an instruction to perform a particular comprehension operation. If perfective marking increments the model to a new situation, progressive marking sets a stage for further events, while habitual marking adds the information to the file of information associated with the character (e.g., *John walks through the kitchen into the bedroom after the evening news every night.).*

It appears then that tense-aspect morphology and morphology which relates verbs (and adjectives) to their arguments fit naturally into the picture of grammatical

morphemes as cues with trigger and constrain the construction and modification of mental models or the progression of developing models.

2.2.1.3. Diachrony and inflectional cues

A major purpose of this chapter is to emphasize the powerful role of grammatical morphology in language processing and thus to resist any temptation in SLA studies to view grammatical morphemes a mere formal requirement, a matter of putting the right pieces in the right places during spoken production. Far from being unsystematic redundancies resulting from accidents of historical development, grammatical morphemes are the result of a powerful tendency of languages (Bybee, Perkins and Pagliuca, 1994; Hopper & Traugott, 1993; Heine, Claudi & Hünnemeyer, 1991) which is best understood against the backdrop of their processing role and the associative nature of form-function learning discussed in the previous chapter.

In the evolution of a form from lexical to grammatical status, there is an increase in the frequency of the form involved (Bybee, et al., 1994). If forms are tied to functions, then in order for a form to be frequent, the processing events that the form triggers or constrains must be equally frequent. The gradual diachronic shift in the function of a form ought to reflect a shift from triggering or constraining relatively infrequent aspects of comprehension processes to triggering or constraining relatively frequent aspects of comprehension processes. For example, lexical items meaning 'finish' or 'complete' often develop into grammatical markers of completive aspect ('to do something thoroughly and completely', Bybee et al., 1994, p. 57) which in turn develop into markers of perfective aspect (marking events in narratives) or simple past tense. We can see here a movement from a component occasionally present in conceptual representations (i.e., explicit representation of the completion of an activity as a separate step) to a component more frequently present in conceptual representations (that someone carried out an activity in its entirety) to a component that is extremely common in conceptual representations (that an event occurred). In other words, the extremely frequent conceptual property of eventhood (i.e., the effect of moving a narrative forward a step) is ultimately able to preempt the form that was originally associated with the infrequent inclusion of the completion of some activity being a separate step in the narrative.

A comparable development is probably behind the function of Russian verbal prefixes (diachronically deriving from ordinary prepositions) as markers of perfective aspect. A verb which designates an ongoing activity without natural endpoints (*atelic*, Brecht, 1985) can be modified by a one of many prefixes. For example, a prefix can add a notion like "inception, conclusion, intensification, a limited period of its duration, or the like" (Brecht, 1985, p. 15). Thus the generally atelic transitive verb *kurit'*, 'smoke (cigarette, etc.)' can be modified by prefixed *vy-*, in which case it carries the idea of finishing smoking. The prefix *za-*, adds the idea of starting smoking. The prefix *na-*, adds the idea of smoking to satiation. Other prefixes are less obviously aspectual in their semantics. For example, the addition of *pere-* to a verb can give the notion of overdoing the action, or redoing the action. The fact is that all such prefixes (whether sub-lexical or lexical) have the effect of creating verbs that describes events that moves the discourse to a next stage. The relationship of prefix to event status became so strong that a prefix must now be added to most unprefixed verb stems simply for the sake of marking eventhood (i.e., marking perfective aspect). The prefix used differs from verb to verb. The general formal category *prefix* at some point in time came to be the marker of perfective aspect for most verbs. That is, prefixes on the formal level and eventhood on the conceptual level came to be associated due to their frequent co-occurrence. Note that the properties of conceptual representations which can gain grammatical expression must exist on the

conceptual plane in order for those expressions to develop. The diachronic mechanism of grammaticalization, on this account, would appear to be essentially the same as the associative mechanism of learning proposed in Chapter 1.

2.2.1.4. Multifunctionality of grammatical markings

It was suggested above that both case morphology and tense-aspect morphology appear to have functions at the level of the construction of single scenes in mental models and at the level of sequencing the single scenes into evolving discourse models, that is, managing their evolution and relating later parts to earlier parts. The multifunctionality of grammatical morphemes is more far-reaching than this. For one thing, there is a strong tendency for groups of grammatical meanings to be signaled by single inflectional formatives (cumulative exponence—Matthews, 1991). This reflects the fact that particular combinations of grammatical meanings occur together very frequently, making them liable to merger. In our terms, processing operations of the sorts signaled by such inflectional forms also co-occur frequently in specific clusters. For example, having a collectivity (plural) in the role of recipient (dative) is a frequent enough and distinctive enough property of events to make possible the diachronic development of a single marker of dative-plural nouns, as in the case of Russian nouns in *–am* (*mal'chik-am*, 'person-dative:plural').

Another sense in which inflections can be multifunctional is pointed out by Friederici, Wessels, Emmorey and Bellugi (1992), who found that German agrammatic patients were sensitive to inflectional forms as they functioned to mark grammatical categories, but not sensitive to the same forms in their roles as markers of person and number. The study involved two experiments, in both of which the task was word-monitoring. The target word would occur either after a properly inflected word-form or after an improperly inflected one. Healthy control participants showed a slowed response time when the target word occurred after an improperly inflected verb. In the first experiment, improper inflections altered the grammatical category of the inflected word, either from noun to verb or from verb to noun. This was possible because of the existence in German of noun-verb pairs with a common stem. For example, in place of *ein tanz*, 'a dance', a past tense suffix might be attached to the stem, yielding **ein tanzte*, 'a danced'. It turned out that the agrammatic participants in the first experiment were as sensitive to such errors as were the healthy participants.

In Friederici et al.'s second experiment, the improper inflections did not alter the grammatical category of the inflected word. An example is **er tanztest*, 'he danced', where the verb *tanztest* has the second person singular ending but the pronoun subject is third person singular. In this experiment, the agrammatic participants showed no effect of the ungrammatical forms. This is reminiscent of Luria's Russian patients discussed earlier for whom case endings appeared to be inert, that is, appeared not to function as processing triggers. However, taking Friederici et al.'s two experiments together, it appears that for those German agrammatics the inflections are not entirely inert. They are apparently active as cues to the grammatical category of the inflected word (noun or verb), but not as cues to the person value of the subject (second or third), having become largely inert in relation to the latter function.[12] We can think of such an inflection as "containing" two instructions (perhaps among others). One instruction relates to employing words as nouns and verbs and the other constrains the merger of a verb-related concept with an argument concept. The damaged processor is able to carry out the

[12] In processing terms, lexical categories may not be theoretical primitives. Being a noun or verb might be reducible to being a trigger of certain types of processes in mental model construction.

first instruction, but not the second. The inability of a listener to make use of the agreement cue does not mean that s/he will be unable to identify the subject of the sentence, since other cues may still be active (e.g., the subject pronoun).

This probably does not exhaust the functions served by the inflectional forms examined by Friederici et al. For example, inflectional endings (or prefixes) could have a demarcative function (helping the processor to find word boundaries), or play other roles in word recognition.

It may be, then, that there is no such thing as a grammatical morpheme with a single processing function. The types of multifunctionality discussed so far involve simultaneous (and hence compatible) functions of grammatical morphemes. From the learner's standpoint, simultaneous, compatible functions of grammatical morphemes might cause less difficulty than situations where multiple functions of a form are incompatible (that is, cannot be simultaneously present). One of these is straightforward inflectional homonymy. For example, the English ending *–s*, may be an agreement marker or a plural suffix (*John likes eggs*), a possessive marker (John's eggs), and may also be a reduced form of auxiliary *has*, or of auxiliary or copular *is (John's here; John's been here*).

Besides such inflectional homonymy, there is what might better be labeled inflectional polysemy. Traditionally, this is described in terms of a single inflectional category having a variety of uses. For example, the verbal *–s* suffix in English (third person agreement) is commonly used to express habitual aspect (*My mother bowls on Thursday evenings*) and permanent states or properties (*My mother likes to bowl*). It is also used for foreground events, use traditionally referred to as the *historical present* (*I was standing in line when this guy reaches out and shoves me*) and in narrating movie plots (heard frequently, for example, in film reviews on television).

Both inflectional homonymy and inflectional polysemy are well represented in Russian and may affect the rate of acquisition. This will be discussed further below.

2.2.2. The formal expression of inflectional cues

The acquisition of inflection involves the association of the formal inflectional cues with the processes they trigger. We have seen that a single inflectional cue may trigger different processes on different occasions, or on a single occasion, a single inflectional cue may trigger several compatible processes. A question that has not been adequately addressed yet is that of the form that "a single inflectional cue" might take. The short answer is that an inflectional cue formally consists in the modulation of word form. That is, to process an inflected word, the listener must exploit the word's basic lexical meaning (i.e., add a component or property to a mental model) and also execute other processes (thereby building the lexical meaning into a complete model and moving an evolving model forward). The processes triggered by a word-form above and beyond the lexical process are triggered by the fact that the word has the particular form it has and not some other form that it might have had. However, to say that an inflectional cue involves a difference in word form leaves much still needing to be said. What exactly are the elements of word-form differentiation? Inflectional form gives rise to its own levels of complexity, and to what may appear to be autonomous linguistic categories that are independent of the level of phonetic form and the level of conceptual form.

2.2.2.1. Bound forms vs. function words

Allowing that an inflectional form can have multiple functions, it might appear desirable that that inflectional form itself have a uniform expression, consisting in a

continuous stretch of sound, such as a perceptually salient prefix or suffix. This would mean that in terms of complexity of formal expression, bound grammatical morphemes would be processed in a manner similar to many function words, which are also constituted by recurrent, continuous stretches of sound. This would appear to be true of some bound morphology in some languages. In a polysynthetic language such as Blackfoot, many sub-word units are reminiscent of function words. Consider (3):

(3) Kit-ak-omat-oxkott-oxkana-iksist-anist-aax-i't-aki-atts-ootsp-oaawa
you-will-begin-able-all-equal-manner-good-feel-cause-relator-you:plural
An effort will be made to satisfy everyone equally.

It seems unlikely that such words could be understood in any way other than by many relevant stretches of sound being reacted to as processing triggers. The challenge of processing such polysynthetic words could be simplified if many reasonably frequent combinations of two or more morphemes are stored as units in the mental lexicon. However, it seems unlikely that the processing of such a word such as (3), which takes approximately four seconds to utter at a normal speech rate, would be postponed until all twenty-one syllables have been perceived by the listener. Thus, in principle, bound grammatical morphemes, including those expressing classical grammatical meanings, could be realized by sub-word sound stretches which stand in a one-to-one relationship, if not with single processing functions, at least with sets of simultaneously triggered processing functions. It is reasonable to think that in agglutinative languages (defined loosely as languages in which there are more nearly one-to-one form-function relationships for inflectional affixes) affixes could be processed like function words. In that case, the statement of the form of inflectional cues would be relatively straightforward. Identifying an inflectional cue would be in principle similar to identifying a word.

2.2.2.2. More complex cues: beyond simple concatenation

In many cases the formal expression of inflectional cues is considerably more complicated than this, and it becomes difficult to think of bound morphemes as the word internal analogue of uninflected function words. In addition to cumulative exponence of inflectional categories (discussed above), inflectional expression can involve what Matthews (1991) calls extended exponence. The marker of perfect aspect in English, for example, is divided between two nonadjacent locations (boldface): ***I have taken it.***[13]

In general, the modulation of word form for grammatical purposes can take a variety of forms besides the simple concatenation of stems and affixes (Hockett 1947, 1954; Nida 1948, 1949; Matthews, 1991; McCarthy, 1981; Anderson, 1992). These include infixes, circumfixes, zero morphs, reduplication, reductions, vowel alternations, consonant alternations, stress changes, tone alterations and/or suppletion, with many of these possibly participating in distributed exponence or cumulative exponence. In

[13] Note that *have* here also illustrates cumulative exponence, since besides contributing to marking perfect aspect, it also differs from *has* and *had* along other inflectional parameters. Besides that, the word-form *have* has another use as a possessive verb. Thus the expression of perfect aspect in English, besides being conceptually complex, as discussed earlier, is also formally complex. There is an entangled in a many-to-many form-function relationship which involves extended, discontinuous exponence. From such considerations, no doubt among others, we might plausibly predict that learners of English would require considerable experience with the language in order to develop the form-function associations that constitute the English perfect aspect system.

American Structuralist linguistics (e.g., Bloomfield, 1933), inflectional affixes were treated as lexical entries, concatenated to stems to form words, just as words are concatenated to other words to form phrases. Today a few morphologists (e.g., Lieber, 1992) treat inflectional affixes as straightforward lexical entries that are joined to stems by syntactic concatenation. It is perhaps more common to treat inflection as consisting in processes of word modification (e.g. Anderson, 1992; Beard 1995), or to locate the various word-forms of a lexeme in the lexicon, organized around a base-form (Bybee, 1985; Günther, 1988; and from a psycholinguistic perspective, Lukatela et al., 1980; Jerema & Kehayia, 1992).

2.2.2.3. More complex cues: forms-in-relation-to-other-forms

A strong case can be made that at least in the case of Russian, the inflectional values expressed by a particular word-form can only be understood in relation to other forms of the same lexeme. In many cases, this would appear to require that at least one already inflected form, the base-form, be stored in the mental lexicon. Zakharova (1973 [1958]) claims that in connection with learning L1 Russian,

> Experiments show that in order to produce forms correctly, it is important for a child to assimilate the structure of the word [noun] in the nominative. (Zakharova, 1973 [1958], p. 283)

Even for English, a weak case can be made for considering an inflectional cue to be not the mere presence of a word ending, but the presence of that ending in contrast with its absence in the base-form. Consider the inflected word *paws* (as a plural noun). What indicates that it is plural? We cannot simply say that the final *–s* indicates that the noun is plural, since the word *pause* ends in essentially the same sound (at least the sets of possible pronunciations of *paws* and *pause* probably intersect). Therefore, we might need to say as a minimum that an inflectional modification serves as a processing cue (has a "meaning") by virtue of the relationship of the inflected word to other forms of the same lexeme. That is, *-s* counts as an instruction to activate a collectivity token in the mental model in the case of *paws*, because there is a singular form *paw*. In understanding *paws*, in some sense there must always be implicit reference to the form *paw*.

This may appear trivial in the case of English plural inflection. In the case of Russian nominal inflection this lexical-relational nature of inflectional forms is more striking. The ending *–u*, for example, can be a marker of the inflectional categories *accusative, dative, instrumental, locative* and *partitive,* depending on the base-form. Generalizations regarding when *–u* marks which category can only ultimately be stated in relational terms: *-u* marks the accusative category if the lexeme has a word-form in *–a* as the representative of the nominative category; *-u* signals the dative category if the lexeme has a word form in *–Ø* signaling the nominative category; etc. In fact, the lexical-relational statements might need to be more complicated, at times taking into account more than one other word-form in the paradigm of a lexeme. From the standpoint of lexical access during comprehension, we saw evidence in Chapter 1 that in hearing a word-form such as *mal'chik-u,* "boy-dative", the base-form *mal'chik,* "boy-nominative", might be activated as part of the cohort activation process, since *mal'chik* is the initial portion of *mal'chika.* On the same basis, during the processing of *zhenshchin-y,* 'woman-accusative', the base-form *zhenshchin-a* 'woman: nominative' and other word-forms in the same paradigm, might also be activated. That is, the relating of a particular inflected form to other members of its paradigm might happen readily during native language processing. Nevertheless, the lexical-relational nature of inflectional forms complicates our

characterization of the formal expression of inflectional cues and affects our understanding of what is being acquired if or when L2 inflection is being acquired.

2.2.2.4. The emergence of "autonomous" morphological form

There is another, perhaps more far-reaching, sense in which inflectional form can only be understood in relational terms. Consider again the morphology of perfect aspect in English. In the phrase, *has eaten*, the perfect aspect is marked by the auxiliary *has* and the suffix *–en*. The first part of this discontinuous marker, *has*, employs not a single particle, but of a whole paradigm, which is, in fact, identical to the paradigm of the lexeme *have*, a verb expressing the meaning of possession. It is as though that paradigm is picked up, lock, stock and barrel and redeployed as a cue to perfect aspect (a phenomenon sometimes dealt with in morphological theory by *rules of referral*; see Zwicky, 1985 and Stump, 1993). It may be less obvious, but the same sort of lock-stock-and-barrel reuse of an entire cue network is involved in the word form *eaten* of the phrase *has eaten*. In traditional terminology, the form that is employed here is the *past participle* form of the lexeme *eat*. The participle form is the form used for deverbal adjectives, as in the phrase *already eaten food*. The past participle forms of *eat, get, put, sing, bring* and *bake* are *eaten, gotten, put, sung, brought and baked*, respectively. That is, there is a fair bit of irregularity involved. This makes it easy to see when this entire cue network is redeployed. This happens not just in the case of the perfect construction, but also with the verb of the passive construction, as in *was eaten, was brought, was sung*, etc. Thus the "cue" that marks the main verb in the phrase, *has eaten* is not merely the suffix *–en* (and not merely the form *eaten* contrasted with *eat*, as discussed in the previous section), but rather the suffix *–en* (or form *eaten*) as a member of a whole network of inflectional modifications (including those seen in *eaten, gotten, put, sung, brought and bake*). Such reusability gives the network of past perfect forms an appearance of autonomy. This apparently autonomous level of morphological form is called *morphomic* by Aronoff (1994), a level which is "neither morphosyntactic nor morphophonological, but rather purely morphological—morphology by itself" (Aronoff 1994, p. 25).

In processing terms, membership in the *–en* network appears be a property of a word form which the processor can make use of. Call this property *P*. To have property P is to be a part of the network of word-forms *eaten, gotten, sung, brought, baked, etc.*, that is, part of the set of words that correspond to the aspects of processing triggered by the members of the inflectional set in its different functions. A child's initial acquisition of property P need not require an autonomous morphomic level. As soon as this set of forms has taken on a single function, the property P exists. It is nothing more or less than the property of being one of the forms in the network associated with some function.

For example, take the function of the *–en* set when it marks the verbs of "get passive" sentences such as *I got pushed*. By virtue of its association with this function, the verb form *pushed* is associated with all other verb word-forms which are also associated with this function. This is an essential aspect of Bybee's (1985) lexical model. In the present connection it gives concrete substance to the notion of property P. When the child has learned the set of verb word-forms used in this passive construction and later on is confronted with the same set of word-forms in the perfect construction, the property P can be exploited and the set of forms carried over lock, stock and barrel. At that point it appears that the cue that is being exploited for the new function is precisely the property P, that is, the property of belonging to a network of forms related in a common way to an existing processing function. If such a cue is reused for a number of functions, then it will appear to have more and more autonomy with each function that is added, giving credibility to the claims of autonomous morphology. As will be noted later,

it is this level of morphological cues that appears to be involved in the notion of case cues employed in the Competition Model (at least in Kempe & MacWhinney, 1998).

In short, multiple uses of networks of forms in fully formed adult (native) language systems, such as the multiple uses of the *–en* network, need to be understood against the backdrop of L1 acquisition. That is, in looking at the full-blown adult system, we see autonomous morphology.[14] But it need not have been acquired (and perhaps could not have been) as such at the outset. Rather there appears to be the pattern formulated by Slobin: "New forms first express old functions and new functions are first expressed by old forms" (Slobin, 1973, p. 184). In the case of the property P (=membership in the *–en* network), its association with perfect aspect is one of the latest developments in the acquisition of English inflection, occurring at the age of four or five (Clark, 1998), some time after the *–en* network is established for earlier functions. Property P is then recognized as the cue for the new processing function, even though the new function is conceptually unrelated to earlier functions. From the child L1 learner's standpoint, given that a particular processing function is going to be cued by irregular morphology, reusing an already acquired set of forms ought to be easier than acquiring a new set. From the adult L2 learner's standpoint this may not be the case. Rather, the adult learner may be under pressure to deal with multiple functions from an early point, in which case the existence of many-to-many form-function relationships could hinder the formation of associations, rather than facilitating it. In the case of L2 Russian, the magnitude of this problem could be great indeed.

2.2.3. Summary

Inflectional cues would potentially appear to be one of the most complex aspects of language, both in terms the comprehension processes they trigger and in terms of their formal expression. In the cascade of cues and corresponding processes, going from acoustic cues to mental models, it appears that in some languages there is a level at which membership in a particular network, defined by the original association of word-forms in that network with their earliest acquired function(s), serves as a cue to further processing. To identify a word-form as a member of that network involves reference to other forms of the lexeme to which that word-form belongs. In certain functions, that cue may be part of a larger, discontinuous set of cues (extended exponence). A single inflectional cue may trigger several compatible processes. A single phonetic form (e.g., a suffix) may belong to more than one cue. Add to all of this the fact that inflectional cues are tightly united with lexical cues and that a lexical cue alone will exact a clear processing cost at almost the same time that the processor must deal with the inflectional cue.

Granted, from a processing perspective all of the complexity related to an inflected form need not play a role every time the inflected form is encountered. When an inflected form is encountered for the first time there might be a need to make reference to other forms of the same lexeme, to the membership of that inflection in a set of inflected forms and so on. However, once a given form has repeatedly triggered or constrained a group of comprehension processes, the association between a specific inflected word and the comprehension processes it triggers could become direct. When it comes to forming associations between cues and processes (=learning) in the first place, it is harder to see how the complexity can be avoided.

---

[14] See Beard 1985 for an elegant sketch of the sort of complex algorithm that would appear to be necessary to map grammatical roles—already a step removed from semantics—onto specific aspectual allomorphs.

2.3. SELECTED ASPECTS OF RUSSIAN INFLECTION

Inflectional modulation of Russian words is pervasive. It affects nouns, pronouns, adjectives, verbs and various closed-class categories. In the experiments reported in the next chapter, I am primarily concerned with case inflection of nouns and aspectual inflection of verbs, with side glances at verbal person agreement and gender agreement. Therefore, in this section I emphasize case inflection and aspect inflection.

2.3.1. The form of Russian case cues

Russian case-marking was already cited above in support of the relational character of some inflectional cues. Consider the word-form *mashin-u* 'car-accusative', as seen in the utterance *Mashinu poveslo* 'The car skidded' (literally, 'It.impersonal carried the car'). As noted, the superficial view of the formal cue as consisting of the ending *–u*, 'accusative' is inadequate. The ending *–u* by itself can mark a variety of cases. We noted the need for a formulation such as the following informal one: *–u* is a marker of accusative case by virtue of its occurrence on a noun which has a nominative case form ending in *–a*. (Of course, it is question begging to throw the phrase "nominative singular form" into this definition of the cue, since we are appealing to another cue that we have not yet characterized and that raises the same problems.)

The characterization in terms of the base-form in *–a* is not entirely adequate due to phonological neutralization of /a/ and /o/ in unstressed syllables. Phonetically, *derivo*, 'tree' and *mashina*, 'car', both end identically, the ending in both cases being [ə]. So to characterize the accusative *–u* by reference to nominative forms in -[ə] is inadequate, since the word-form *derevu* is not the accusative form of *derevo*, 'tree', but rather the dative form. We might say that the cue to accusativity, in the case of *mashinu*, 'car:accusative', is "the ending *-u* on a noun whose nominative ends in -[ə] and whose dative form ends in -[ɪ]". That might work, as might reference to other ad hoc combinations of parts of the paradigm.

Note that the ending -[ɪ], 'dative', raises similar problems. The endings of the orthographic forms *n'an'a*, 'nanny:nominative', *n'ane*, ambiguously 'nanny:prepostional.case' or 'nanny:dative' and *n'ani*, ambiguously 'nanny:genitive:singular' and 'nanny:nominative:plural' are phonetically identical in standard Russian (Bondarko, 1998). (According to Bondarko, native speakers might debate the reality of this phonetic neutralization. However, from the standpoint of articulation and speech perception, he holds it to be empirically established [1998, p. 179]).

The endings –u, -[ə] and -[ɪ] illustrate the broader issue of case-neutralization. Shvedova et al. (1982) point out that in every declension there is case neutralization somewhere in the paradigm, even without phonetic reduction. For example, the distinction between nominative singular and accusative singular is morphologically marked only for second declension nouns and animate masculine first declension nouns. (Perhaps not surprisingly in the face of so much neutralization, Offord, 1996, reports northern dialects of Russian for which the nominative-accusative case distinction is no longer marked at all.)

Other neutralizations include the genitive singular and accusative singular of animate masculine nouns of the first declension (e.g. *muzha*, 'husband:accusative' or 'husband:genitive'); the dative and prepositional cases of second declension nouns (e.g. *mashine*, 'car:prepositional' or 'car:dative'); and the dative, prepositional and genitive

cases of third declension nouns (e.g., *dveri*, 'door:dative', 'door:prepositional' or 'door:genitive').

In terms of language processing, an important question is what happens when case-neutralized forms are encountered. In Chapter 1, we saw that the preponderance of evidence points to a difference between the handling of lexical ambiguity and syntactic ambiguity during comprehension. In the case of lexical ambiguity, it appears that multiple homophones are activated simultaneously. In the case of syntactic ambiguity, it appears that more commonly the parser performs a single initial parse. Ambiguity of case marking could be analogous to syntactic ambiguity or to lexical ambiguity. In addition, it might follow a third alternative: not dealing with case at all until additional cues settle the ambiguity. Thus on encountering utterance initial *loshad'*, 'horse', where the distinction between nominative and accusative case is neutralized, the parser could rule out oblique cases, but simply remain uncommitted on the nominative-accusative dimension until the issue can be decided later in the utterance. Of the three alternatives: (1) activate all neutralized case values, (2) choose a single, "most likely" case value and (3) ignore case to the extent that it is not distinguished. Kempe and MacWhinney (1999) conclude that Russian listeners (and German listeners, to a smaller extent) follow the second alternative. A noun that could be taken as either nominative or accusative is initially taken as nominative. This can turn out to be a garden path, requiring a reanalysis. In other words, Kempe and MacWhinney's evidence suggests that case-marking ambiguities behave more like syntactic ambiguities than like lexical ambiguities.

Kempe and MacWhinney (1998, 1999) do not discuss the question of *what* constitutes the case cue. Rather, they simply take a case value in the traditional sense to be the cue, however that case value might itself be cued (i.e., physically expressed). If an utterance initial noun is unambiguously nominative, regardless of how that fact is determined, that fact itself is the cue to subjecthood (or agenthood). Likewise, unambiguous accusative marking is a cue to objecthood (or patienthood). The case cues would thus appear to exist on the morphomic level discussed above. That is, the case cue has to consist in the membership of the case-marked word-form in a network of word-forms belonging to different lexemes and united by the shared processes they trigger. Such an understanding of case cues avoids circularity if (1) there is a straight-forward form-function association which can be developed initially and (2) the set of word-forms resulting from that form-function association can be extended to new functions.

2.3.2. Some functions of Russian case cues

It is traditional in discussions of Russian (and other languages with multiple functions of case categories) to enumerate a set of functions filled by each case (see, e.g., Shvedova et al., 1982; Koctomarov & Mitrofanova, 1988; Wade, 1992). In modern linguistics, there have also been efforts in the opposite direction, that is, efforts to find single invariant functions for each of the values of inflectional categories, or at least for the marked values (since unmarked values can have "elsewhere" functions). This direction in the analysis of Russian is largely attributed to the influence of Roman Jakobson (see Kilby, 1986; Timberlake, 1982 for discussion). From a contemporary perspective, such a strategy is doubtful. From the standpoint of child language acquisition, Slobin has repeatedly emphasized (1973, 1982a,b, 1985; Berman & Slobin, 1994) that the functions of particular forms are often added to the system separately, over considerable periods of time. In addition, from the standpoint of diachronic grammaticalization processes, the development of multiple (non-invariant) functions is a standard stage on the path of evolution of grammatical morphemes from earlier content words (Bybee, Perkins & Pagliuca, 1994).

In many instances, invariants of grammatical meaning are difficult to conceive. The central function of accusative case in Russian is to mark the direct object, most often a semantic patient or theme, as in (4).

(4) Mashin-u povezlo
Car-acc carried:neuter.subject
The car skidded (literally: [Unspecified.subject] carried the car).

In processing terms, we have taken the position that direct objecthood triggers or constrains (or perhaps confirms) the merger of two concepts, one related to the verb and the other related to the object being affected by the action. In the case of (4), the verb requires a direct object selectionally restricted to entities that skid. The car thus satisfies the argument structure of the verb and the selectional requirements. The case-marking converges with the lexical and semantic cues. The result is a mental model in which a car skids.

A second use of accusative case in Russian is to mark the *destination* of the movement of a theme or of a moving agent:

(5) Ja poshol v komnat-u
I go:past:masc:sg into room-accusative.
I went into the room.

We can imagine an extension of patienthood to destinations such as the room in (5). It might be argued that the room is affected, or changed, by something going into it.

However, consider another function of accusative case: repetition (Wade, 1992).

(6) On eto govoril tys'ach-u ras.
He this said:impf thousand-acc times
He said this a thousand times.

Or consider the use of accusative case following the preposition *s*. The preposition *s* is ambiguous. Its primary meanings are 'off of', in which case its complement nominal is in the genitive case and 'with' (in the sense of accompaniment), in which case its nominal complement is in the instrumental case. When its nominal complement is in the accusative case, it has a comparative meaning, as in (7):

(7) Sobak-a razmer-om s korov-u.
Dog-nom size-instr preposition cow-acc
A dog the size of a cow.

(This use of the accusative requires that the nominal expressing the scale of comparison [size, height, age] be in the instrumental case.)

At this point, any relationship to the accusative of direct object would appear to be difficult to justify. Each case category is traditionally held to have a small number of primary uses and a range of minor uses. Minor uses of the accusative are seen in (6) and (7). The use illustrated in (5) would be considered a major use.

In addition to the sort of multifunctionality illustrated in these examples, there are other possible functions of case marking of the sort discussed above in the general discussion of multifunctionality. For example, a case ending might aid in the rapid determination of word boundaries and in determining the lexical category.

2.3.3. How can such complex form-function relationships be learned?

We might expect the many-to-many nature of Russian inflectional form-function relationships to make it difficult for learners to develop automatic responses to inflectional cues. In much simpler tasks Schiffren and Schneider (1977; Schneider & Schiffren, 1977) found that the development of automatic responses was hindered when there was not a "consistent mapping" of stimulus sets to responses. With Russian case-marking, we saw that a single ending can be used to mark various case values. Conversely, each case value can have various formal markings, depending on the noun declension. That is, nominative can have the endings *-Ø, -a, -o*; accusative can have the endings *-Ø,-a,-y,-o*; and so on. In addition, we have just seen that on the conceptual level the possible effects of each case are multiple. We have hardly hinted at the full degree of complexity. The reader might be justified in asking how cue-response relationships can be learned in the presence of such complexity? How can associations form when "what associates with what" involves many-to-many relationships on various levels? And how could associations formed in such circumstances lead to the automatic, high-speed cue-response pairs that are required for language processing in real time?

An important part of the answer to these questions would appear to lie in the principle that new forms are first used with old functions and new functions first expressed by old forms (Slobin, 1973). For a child language learner to get a foothold, imagine s/he were to associate a single case form with a single case function. Suppose that for whatever reasons the *–om* (first declension) instrumental ending managed to become associated with the meaning 'instrument used to perform an action'. At this point, the child's system is "aware" first of all that 'instrument used to perform an action' is a conceptual category that can be expressed linguistically. (Perhaps more accurately, the suffix *–om* has come to be associated with the addition of a token into a mental model in a particular type of relationship to the other tokens and the activity in the model.) Having this conceptual category now available for formal expression opens the way to associating it with the other possible instrumental markings (second declension *–oj*, third declension *–u*). In addition, the relationship *–om* ↔ *processing-instruction* will have given morphological status to the phonetic substance of *–om*, making it available for later morphologically autonomous deployment for other functions. As the other instrumental forms (second and third declension) come to be associated with the same conceptual manipulation, by virtue of that association they will be associated with one another, a fact reinforced by their occasional coordination as in, *shest-om, bit-oj i trostj-u, 'with a pole, a bat and a cane'* (all marked for instrumental case, with the endings appropriate to the first, second and third declensions, respectively). Belonging to the set of forms which are associated with the same processing instruction (in traditional terms, the same grammatical meaning) is then a property which connects every word-form in the whole set of word forms belonging to the instrumental case (Bybee, 1985). The morphologically autonomous property thus emerging can subsequently be used as a cue to other functions, such as marking passive agents, or predicate nominals (later acquired uses of the instrumental case form in Russian).

Such a scenario intuitively makes the learning of such many-to-many relationships involved in Russian case morphology seem more reasonable. But is there any evidence for such a pattern in Russian child language? In fact, Slobin's concept of the addition of new functions to old forms appears to have been influenced at the outset by his study of Russian child language scholarship (Slobin, 1966a,b). Citing Gvozdev (1949), Slobin (1966b, 1982b) discusses acquisition of the instrumental case and the accusative case in the general spirit of the scenario sketched above. For instrumental case

marking, the child began using the first declension suffix *–om* with nouns of any adult declension. That is, his language processor established that a particular concept is formally represented—step one. Unlike my purer scenario, this child appears not to have limited the association to a single function, using it to express the semantic roles of *instrument, accompaniment* and an *object the obtainment of which is the purpose of the agent's movement.* (In adult Russian, the latter two require prepositions in addition to the case, but this child relied on case alone at that point in time.) The fact that all three uses emerged together in the child's speech does not necessarily mean that the associations emerged simultaneously in the comprehension system, of course. For accusative case marking, the ending *–u* was acquired (second declension) and also used with stems belonging to adult declensions requiring other endings. The initial limitation of this suffix to a single function was clear: it was only used to mark objects which underwent direct physical actions carried out by the subject. Thus in the Russian translation of *I tore the book*, the word for 'book' would be accusative marked, while in *I saw the book* it would not be accusative marked, since the book is not affected in any way by being looked at.

The challenge for a child to learn case following such a scenario might not be as great as the challenge facing an adult L2 learner, since the child's universe of discourse starts out small, and discourse containing advanced uses of a particular inflection might commonly be inaccessible to a child at earlier stages. Nevertheless in a language with as much complexity in its case system as we see in Russian, the complete system is acquired late, with errors occurring as late as five years of age (Clark, 1998). If the child, relieved of the early pressure to cope with the ultimate, adult Russian many-to-many form-function relationships, nevertheless requires several years to acquire the case system, then we might plausibly predict acquisition to take longer for adult learners, since they are often forced to cope with multiple forms and functions almost from the outset.

2.3.4. *Russian aspect inflection: function*

The primary aspectual distinction in Russian is between perfective and imperfective. As with case, there has been an effort to find invariant meanings/functions for the Russian aspects (Timberlake, 1982). In Russian scholarship, according to Bondarko (1997), the primary meanings are associated with totality (viewing an action as a whole) and boundedness (viewing an action with its end-points included). Imperfective aspect has been treated as the aspectual "elsewhere".

It is common to relate aspect primarily to temporality, along with tense (Comrie, 1976; Timberlake 1985; Klein, 1995). An alternative emphasis already discussed ties verbal aspect to the sequencing of events in discourse and the foregrounding and backgrounding of those events (Hopper, 1979; Chvany, 1985). Earlier we questioned why such a large variety of languages should develop obligatory devices that would function primarily to remind the listener that every displaced event in a situation being described occurred prior to the time of the utterance. When temporal concepts need to be brought to the listener's attention, that is the natural domain of lexical items with temporal meanings (*yesterday, after, while, next Thursday, awhile ago*, etc.). Obligatory grammatical morphemes, I have argued, are associated with the dynamics of mental model formation and the evolution of mental models through discourses. Clearly, indicating that an event moves the narrative forward a step, bringing about a new state of affairs, would appear to be a more justifiable role for obligatory tense or aspect marking than the role of conveying the temporal relation of the event to the current moment, or redundantly conveying the lexical aspectual semantics of the particular inflected verb stem.

According to Pupynin (1997), Russian children in their early verb use do not make tense-aspect contrasts, so that early reports of events use present imperfective

forms. That is, forms that would first have been used to describe observable, on-going activities are subsequently used to describe past events. When past-perfective forms begin to be used, they are used to describe events whose result is evident in the present observable situation. Later they are used to mark displaced (narrative) events, that is, actions which move the narrative forward a step, without essential concern for how long a step may have taken. However, as the child goes on to manage more textured discourse, we can assume from the pattern of fully developed Russian that perfective marking is not extended to non-event marking functions such as habitual actions or states, or backgrounding functions (with strategic exceptions—see Chvany, 1985).

The picture would appear to be somewhat different for English L1 learning. English also has an early event marker, the simple past tense. At a certain stage, the English simple past and Russian past perfective could be functionally similar, being used to mark displaced events. English and Russian would subsequently differ in the pattern according to which new functions come to be added to old forms. The (relatively) early Russian marker of displaced events (perfective aspect) continues to be reserved largely for that function, while the early English marker of displaced events gets extended to non-event functions as in (8) and (9).

(8) While I play**ed** in the park, my brother tried to make his way home.

(9) When I was a kid in Toronto I often play**ed** in the park.

In Russian analogous sentences which do not advance the event chain would not be marked with the form that is used to mark narrative events (the past perfective form). This difference between English and Russian is commonly attributed to the fact that in Russian aspect is marked in contexts where English marks only tense. An alternative view of the difference, however, is that English and Russian differ in the pattern of extending early forms to later functions. The early English event marking morphology gets extended to certain non-event marking functions while the (relatively) early Russian event marking morphology does not. This means that in Russian the form of imperfective aspect is used for a wider range of functions than the form of perfective aspect. The perfective may be used with imperative morphology to request the execution of an event, or with present tense morphology to state the intention to perform an event (or to predict an event), in addition to its use with past tense morphology to report displaced events. Given the relatively strict limitation of perfective to events, the wider range of imperfective functions does give imperfective the status of the "elsewhere" aspectual value in Russian. This could conceivably make learning the multiple functions of aspect marking simpler than learning the multiple functions of case marking, where the non-primary uses of cases cannot be viewed as linguistic "elsewheres".

In the broad sweep of comprehension processes, aspectual cues, I argued, relate primarily to the ways clauses are integrated into discourses. An "event" is a feature of a discourse that moves the plot ahead one step. A background activity (Russian past imperfective, English past progressive) contributes to setting the stage for events which occur on that stage (among other functions). A background habitual fact, for example, the fact that a character worked at a particular place (Russian past imperfective again, English simple past) adds information to the file of information attached to an individual in discourse so that that information can be used for inferential processes and contribute to discourse cohesion in various ways. These could be called major functions of the imperfective. There are also minor functions, such as marking an unsuccessful attempt to perform an action (sometimes called the conative function).

2.3.5. Russian aspect inflection: formal expression

Turning from the functions of aspect marking, having again just scratched the surface of an enormously complex topic, let's briefly consider the formal nature of the aspectual cues in Russian. This aspect of Russian is also considerably more complex than may be apparent from the comments below.

2.3.5.1. Verbal prefixes and aspect

We noted above that the meanings expressed by verbal prefixes vary, but have in common the feature that they derive event verbs. Take for example the verb *bezhat'*, 'run' which designates an ongoing activity rather than a brief event. The prefix *ot-* has a basic meaning of 'motion away from a specific place'. Thus *otbezhat'* has the meaning 'run away from a specific place'. However, the act of running away from a specific place, i.e., departing that place by running, is a brief event. This illustrates the general way in which prefixal modifications of verbs add some specific meaning but in the process convert activities into events, so to speak. In the case of *otbezhat'*, 'run away', the word formation process (adding the prefix to the stem) can be considered productive: any verb designating some type of motion can be combined with the prefix, and the result will be transparently semantically compositional. In other cases, a prefix has become fossilized as part of a stem so that the combination of prefix and stem is semantically idiosyncratic. For example the stem *kryt'* has the meaning 'cover'. With the prefix *ot-* however, it forms a semantically shifted lexical item, *otkryt'*, 'open'. Such semantically shifted prefixed verbs have almost never lost the grammatical property of perfectivity, unless further derivationally marked as imperfective, and their lexical semantics is generally compatible more with events than with processes or activities.

As noted earlier, given that prefixes co-occur reasonably consistently with narrative events, that is, with incrementations of mental models to new states of affairs, this association has taken on a partly independent status in that prefixes often appear to be present solely for the purpose of marking eventhood (perfective aspect). In some cases they do have a partly non-arbitrary relationship to the meaning of the verb involved. For example, the verb *bogatet'* has the meaning 'become rich' and conveys the idea of a process, not an event. In a narrative context, becoming rich can be treated as an event. That is, there can be a change from one state of affairs to another and the time taken for the change to occur can be totally ignored. In that case, there is an imperfective form, *razbagatet'*, *'become rich'*. The suffix *raz-* is the only way to perfectivize the imperfective *bagatet'* and in a sense then, it is a lexically neutral (i.e., purely grammatical) imperfectivizing prefix. On the other hand, the primary meaning of *raz-*, which has to do with movement outward in multiple directions, seems to have a metaphorical compatibility of becoming rich, involving a metaphor of increased wealth as enlargement. Similarly, *raz-* is the "neutral" perfectivizing prefix for other verbs whose result may involve metaphorical or literal expansion, *zlit'*, 'make angry', *serdit'*, 'make angry', *svirepet'* 'grow violent', *toloch'*, 'grind' and *davit'*, 'crush'. (With some of these stems, the existence of multiple perfective forms corresponding to a single imperfective form makes it difficult to prove that the form in *raz-*, is *the* neutral perfective form.) The point is, to perfectivize such verbs, it is necessary to add a prefix which may not be entirely semantically unrelated to the meaning of the verb, at least on a metaphorical plane. With other verbs, a particular prefix does appear to function as a truly neutral perfectivizing suffix. Thus in the case of *kryt'*, 'cover', the perfective is formed by adding the prefix *po-*, i.e., *pokryt'*, 'cover:perfective'.

In summary, a formal cue to the conceptual property of eventhood is the presence of a prefix (which may or may not contribute to the meaning of the verb in other ways). Shvedova et al. (1982) list twenty-eight verbal prefixes. The relationship of a given prefix to features of the mental model building process can clearly be one-to-many, given the fact that the prefix may contribute its own meaning (such as the spatial-directional meaning of *ot-* 'motion away from a specific place') as well as signaling perfective aspect (i.e., eventhood).

Going in the direction from function to form, devices for representing perfective aspect are numerous: twenty-eight prefixes that can cue perfective aspect, in addition to other means such as a suffix, stress shifts and suppletion. So with perfective aspect, although we don't see the same level of unprincipled homophony that we saw with case endings, where a single suffix marked different cases depending on the declension (*e.g.*, *-a* marking nominative singular in the second declension and genitive singular and animate accusative in the first declension) we do find a complicated many-to-many relationship between form and function. To complicate matters further, the prefix cue to perfectivity is not a reliable cue, in that it is often cancelled out by the addition of an imperfectivizing suffix (referred to as *secondary imperfectivization*).

2.3.5.2. The marking of imperfectivity

Turning then to the formal cues of imperfectivity, with unprefixed verbs, the basic unmarked form is imperfective and a "neutral" prefix must be added to derive the corresponding perfective. The stem *ctroit'*, 'build' designates a process, which can be treated as an event in narrative (i.e., moving the narrative forward a step while completely ignoring the amount of time that the step might have taken in reality) by adding the neutral imperfectivizing *po-*. However, we also saw that the addition of semantically productive prefixes has the (possibly unwanted) side-effect of producing a perfective verb. If a speaker wants to say that someone *rebuilt* something, the prefix *pere-* productively adds the desired meaning, deriving the new word *perestroit'*, 'rebuild'. Having a prefix, this verb marks an event of rebuilding. If it is necessary to talk about a present activity of rebuilding, a background process of rebuilding, or repeated, periodic rebuildings, etc., the perfective morphology of *perestroit'* will be incompatible with such non-event meanings. For such situations there are relatively productive suffixes available to imperfectivize such prefixed verbs. The total statement of the morphological change is complicated in the case of *perestroit'*, but the result is the now imperfective prefixed form *perestraivat'*. The derivation of imperfective stems from prefixed (and hence perfective) stems is very widespread in Russian, usually involving the addition or substitution of a suffix.

As with cues to perfectivity, there is morphological variety in the possible cues to imperfectivity. Rozental', Golub and Telenkova (1995) list the following: (1) no marking (simple imperfectives); (2) suffixes (-iva- ~ -yva-; -va-; -eva-; -a-; and replacement of -i- by -a-), often with accompanying stem consonant changes; (3) stress change; and (4) suppletion. As with case-marking, we end up with a picture of many-to-many form-function (cue-process) relationships. Here too, the problem of establishing the necessary mental associations (that is, the problem of learning) would seem insurmountable if their were not some piecemeal, gradual development scenario in which the first step is the establishment of the fact that certain conceptual events/properties are formally cued, then the discovery of new formal cues related to the same events/properties and then the extension of entire functioning cue-sets to new functions. The final adult grammar might be described by a linguist-analyst as an integrated whole, but this surely obscures the nature of the developments that would allow such a processing system to take shape.

Given the variety of aspect markings, with none clearly predominating, it does not seem likely that children would learn a single marker of perfectivity or imperfectivity and apply it to verbs generally.

A more plausible scenario might be that a child first would learn verbs in a single form, generally imperfective, being the form used to talk about the here-and-now. The child with a stock of familiar imperfective verbs could then start learning their paired perfective counterparts one by one, perhaps particularly when the child starts processing large volumes of narrative (Thomson, 1995). In effect, the child would have two verbal lexicons, an imperfective lexicon and a perfective lexicon. Members of each set would have certain formal similarities, such as the presence of a prefix (any prefix). One set of verbs would be associated at first with ongoing, observed activities and the other with displaced events (or possibly with imperative forms, which tend most often to use the perfective stems). This is basically compatible with the picture of acquisition of Russian aspect presented by Pupynin (1997), and discussed above.

Stoll (1998) investigated the comprehension of Russian aspectual distinctions by native children at five age levels from two to six years of age. With verbs that have straightforward aspectual pairs, that is, with "neutral" perfective and imperfective counterparts, there was no effect of the variable that Stoll called morphology. That is, as the imperfective-perfective contrast was acquired, no particular formal means of expression had any advantage. This too is consistent with the idea that children initially primarily learn many verb pairs. What might the formal cue to aspectual function be at this stage? If the child first generally knows the imperfective form, the formal cue to perfectivity could consist in being "the other form" of an aspectual pair. In most cases, there is a significant amount of formal overlap between the two members of the aspectual pair. Like any formal cue, being associated with a processing function would provide a morphological identity, allowing morphologically autonomous re-use of the cues for additional functions over time. Eventually, more specific cues would become associated with aspect, such as the presence of a prefix (any prefix) or the presence of particular suffixes. But given the complexity of the cues, it seems that the verb-pairing approach would allow a toe-hold and basis for further developments, and the evidence available appears to be compatible with such an account.

2.3.6. Investigating sensitivity to inflectional cues in L2 Russian

In the following chapter, we will be looking for evidence that inflectional cues have become associated with comprehension processes in the language processing systems of nonnative users of Russian. We will examine case cues, aspect cues, and in a limited, way person cues and gender cues. We have seen considerable complexity both in the form of the cues and in the relationship between cues and processes they trigger or constrain. We have also suggested that inevitably the acquisition of such cue systems must be a gradual, time-consuming process. It should be clear now that lexical acquisition and inflectional acquisition are on extremely different orders of complexity. A simple concrete noun, at one extreme, involves linking a time-stable concept to a relatively constant acoustic form. Thus the concept of a dog, rooted in our physical experiences of dogs, can be held in working memory along with a sound-image of the phonetic word *dog* and an association between the sound-image and concept can form. That recurrent bit of acoustic substance will then be a cue instructing the processor to add a dog concept to a developing mental model (or to reactivate a mental dog token that has faded in a narrative model). Consider how different is the challenge of forming the link between the Russian noun suffix *–u* and the processes in the conceptual realm that are associated with it. Therefore, we expect to find in the next chapter that the L2 acquisition of inflectional

cues is a slow process, as is L1 acquisition of these cue systems. In looking at a variety of case cues, aspect cues and person and gender cues we will expect to find that some cue-process associations will for various reasons take hold more readily than others, and strengthen more readily. Therefore, we might predict that levels of sensitivity learners show to various cues will not be equal. That is, that there will be differences reflecting differences in how readily different cue-process associations can be formed. Beyond that, we are not yet in a position to make specific predictions regarding which cue-process associations form more readily than do others.

Before turning to the investigation of sensitivity to inflectional cues in L2 Russian, we need to take a moment to survey some of the major strands of previous research related to the topic of L2 inflectional morphology.

2.4. SLA RESEARCH RELATED TO INFLECTIONAL MORPHOSYNTAX

This section surveys major streams of SLA research related to inflectional morphosyntax. This includes two approaches to grammatical morphology in SLA which have treated it as a unified set of phenomena, as well as work focusing more narrowly on tense/aspect, a particularly active area of SLA research and a tradition which deals with noun case and verbal person agreement as cues to subjecthood (or agenthood). Interestingly, all four of these bodies of work in SLA have received part of their inspiration from research on L1 acquisition.

2.4.1. The Morpheme Studies

The contemporary period of SLA studies is generally traced to seminal articles by Corder (1967) and Selinker (1972). This period is characterized by research into the nature of nonnative language grammars (*interlanguage grammars*) and the ways those grammars change over time. In the early part of this period, there was a good deal of work related to the L2 acquisition of English grammatical morphology. This was a splash-over from research related to L1 acquisition, in which grammatical morphemes were seen as a domain in which growth toward adult-like norms could readily be traced (Brown & Fraser 1963; Fraser, Bellugi & Brown, 1963; Cazden, 1968; Brown, 1973; de Villiers & de Villiers, 1973). Brown and Fraser (1963) asked four children (ranging in age from twenty-five to thirty-one months) to repeat sentences containing various grammatical morphemes. The children primarily repeated the content words stripped of grammatical morphemes. Thus, *I showed you the book* might be imitated as *I show book*. Brown (1973) was able to track the acquisition of fourteen grammatical morphemes by three children and found a consistent order of acquisition, a finding supported in a cross-sectional study of twenty-one children by de Villiers & de Villiers (1973). Although Brown (1973) put considerable effort into attempting to *explain* the order of acquisition (in terms of factors such as frequency, syntactic complexity and semantic complexity), when the investigation of the acquisition of these grammatical morphemes was extended to L2 English, there was less interest in explanation (Wagner-Gough & Hatch, 1975) and more interest in simply demonstrating that there was a consistent order of acquisition across subjects, independent of L1 background, age, learning context and task (Bailey, Madden & Krashen, 1974; Dulay & Burt, 1973, 1974; Larsen Freeman, 1975; Krashen, et al., 1976). Although these claims were challenged on various grounds (Lee, 1981; Pica, 1988; Wode et al., 1978), Larsen-Freeman & Long (1991) conclude:

> ...despite admitted limitations in some areas, the morpheme studies provide strong evidence that ILs [interlanguages] exhibit common

> accuracy/acquisition orders. Contrary to what some critics have alleged, there are in our view too many studies conducted with sufficient methodological rigour and showing sufficiently consistent general findings for the commonalities to be ignored. As the hunter put it, 'there is something moving in the bushes'. (Larsen-Freeman & Long, 1991, p. 92).

One of the major criticisms of the morpheme studies is summarized by Cook (1993):

> On the one hand, the studies usually mix bound and unbound morphemes as if they were the same, despite the correlation found by Krashen *et al.* (1978) between L1 and L2 learners for bound rather than free morphemes. On the other hand, they bring together disparate aspects of grammar. The usual set of nine morphemes includes the morphology of the main verb ("-ing", regular and irregular past tense, third person "-s") the morphology and syntax of the noun phrase (possessive "-s", plural "-s", "the/a") and auxiliary and copula forms of "be"; thus these items blur the conventional linguistic distinction between morphology (grammar below the word) and syntax (grammar above the word) as well as crossing different phrase types. (Cook, 1993, p. 31)

On the other hand, we have seen above that both grammatical morphemes generally, and inflectional marking in particular, do form a natural set of phenomena, in contrast with content words, as proposed by Morrow (1986):

> Whereas content words express object and relation categories (e.g., *car, run*), grammatical morphemes express a relatively small set of conceptual distinctions that apply to most object and relation categories. These distinctions help organize objects and actions into situations, so they must be considered by language users in order to construct a discourse model, a representation of the described situations. Therefore, grammatical morphemes cooperate with content words in order to express situations. (Morrow, 1986, p. 424)

Furthermore, behind the morpheme studies was a recognition that early interlanguage is often short on grammatical morphemes, much like early L1 speech, and that this changes over time. If the pattern of acquisition in L2 development should turn out to be similar to that in L1 development, that would not be a trivial finding. It is also important to keep in mind that whereas L1 learners succeed in acquiring the whole system of grammatical morphemes, for L2 learners this area of language is notable for its tendency *not* to be fully acquired, adding to the potential fascination of this topic.

2.4.2. *Grammatical morphemes as functional heads*

The Principles and Parameters approach to syntax has united grammatical morphemes under the rubric of *functional categories* (Pollock, 1989; Chomsky 1995). In recent versions of this approach, there is a fundamental difference between inflectional forms and function words. Simple invariant function words, which might be, for example, complementizers, modals, negative particles, articles, etc. are simply inserted into syntactic structures in head positions from which they project their own phrases. By contrast, the

functional heads associated with inflection are abstract. They contain features (such as person features or tense features) which do not have phonetic form. Rather inflected lexical items (nouns, verbs, adjectives), are entered into syntactic structures from the lexicon already fully formed, and syntactically move to the positions of these abstract functional heads, where their features and those of the functional head are "checked" against one another. In this way, there is an interaction between inflection and word order, in that inflected words need to move to the relevant Spec positions, either overtly or covertly, for the sake of feature checking.

Whatever the merits of this theoretical distinction between inflectional forms and function words, some scholars have argued that the distinction is at least partly at the root of differences between L1 acquisition orders and L2 acquisition orders as revealed in the morpheme studies. Zobl and Liceras (1994) suggest that the differences can be accounted for by recognizing that whereas in L1 English, within specific categories (noun, verb) inflectional expression is acquired earlier than expression by function words, in L2 English the situation is opposite, with function word expression being acquired earlier than inflectional expression. Specifically, it is claimed that whereas in L1 English possessive marking is acquired before articles are acquired, in L2 English the opposite acquisition order is observed. Similarly, while in L1 English the past tense marking and third-person verb agreement are acquired before auxiliaries, in L2 English the opposite acquisition order occurs. A similar proposal is made for German by Vainikka and Young-Scholten (1998), who suggest that increased difficulty acquiring inflection is a major feature distinguishing L2 learning from L1 learning.

2.4.3. Inflectional categories in the Competition Model

The Competition Model of MacWhinney and Bates (Bates & MacWhinney, 1982, 1989) deals with inflectional categories as being among the (potentially) competing cues that determine the assignment of agenthood and patienthood to nouns in simple sentences. The specific inflectional cues involved in this connection are case marking and verb agreement, cues which can compete or converge not only with one another, but also with word order cues and semantic animacy/inanimacy. This model has been applied to SLA in a number of studies (Bates & MacWhinney, 1981; Gass, 1987; Kilborn & Ito, 1989; Kempe & MacWhinney, 1998). MacWhinney (1997) states in summary that

> We find, uniformly, that the learning of sentence processing cues in a second language is a gradual process. The process begins with L2 cue weight settings that are close to those for L1. Over time, these settings change in the direction of the native speaker's settings for L2. (MacWhinney, 1997, p. 129)

Research in this framework also points to a favoured position for semantic cues (animate beings chosen as agents, inanimate objects as patients) in early SLA if L1 cue-use strategies are ineffective, as summarized by Gass and Selinker (1994):

> ...the research conducted within the Competition Model suggests that learners are indeed faced with conflicts between native language and target language cues and cue strengths. The resolution of these conflicts is such that learners first resort to their NL interpretation strategies, and, upon recongition of the incongruity between TL and NL systems, resort to a universal selection of meaning-based cues as opposed to word-order (or syntax-based) cues. (pp. 142-43)

In connection with Russian case marking, Kempe and MacWhinney (1998) present evidence that case cues are acquired and employed more readily in L2 Russian than in L2 German. Participants in their experiments had been learning German an average of 26.5 months (not counting breaks in learning averaging eight months) and Russian an average of 25.2 months (not counting breaks in learning averaging 9.2 months) (Kempe & MacWhinney, 1996). Only one out of the twenty-two learners of Russian made little or no use of case marking in interpreting OVS (object-verb-subject) sentences, while seven of the twenty-two learners of German made little use of case in interpreting such sentences, despite the fact that the learners of German demonstrated greater familiarity with their target language on the whole than did the learners of Russian ("familiarity" basically being operationalized in terms of vocabulary size). The greater tendency of the learners of Russian to use case cues was attributed to the fact that the case cues are more often unambiguous indicators of subjecthood and objecthood in Russian than in German.

Overall then, SLA studies based on the Competition Model have tended to indicate that over time nonnative speakers do increasingly make use of L2 inflectional cues, that is, SLA does include the gradual acquisition of the L2 cue system, at least in the areas of inflection related to identifying agent and patient (i.e., case and verb agreement).

2.4.4. Tense, aspect and SLA

The acquisition of tense-aspect has come in for considerable attention in the context of SLA studies (Andersen, 1991; Andersen & Shirai, 1994, 1996; Bardovi-Harlig, 1992; 1994, 1998, 1999; Dietrich, Klein & Noya, 1995; Giacalone Ramat & Banfi, 1990; Housen, 1994; Meisel, 1987; Robison, 1990; Shirai & Kurono, 1998; von Stutterheim, 1991). The general consensus is that two factors may influence the early use of tense-aspect, in particular, use of the form typical of event reports, (English simple past, Spanish preterite, Russian past perfective, etc.): 1) the inherent lexical semantics of individual verbs and 2) the status of individual verb tokens in narrative discourse as either foreground or background verb.

The discovery of a relationship in L2 production between the inherent temporal semantics of verbs and the early use of inflections such as the English simple past tense was inspired by similar claims made in connection with L1 acquisition (Antinucci & Miller, 1972; Bloom, Lifter and Hafitz, 1980). The Vendler-Dowty classification of inherent lexical aspectual categories (Dowty, 1979) and modifications of it have generally been used in both the L1 and L2 studies. According to this classification there are four categories of verbs: (1) states (e.g., *know, admire*—referring to conditions that do not require ongoing effort for their continuation); (2) activities (e.g., *play, swim*—processes that can only be continued with the in the presence of ongoing energy input and that have no essential endpoint); (3) accomplishments (e.g., *build, draw*—actions that require an input of energy but also have an essential temporal end-point); (4) achievements (*arrive, fall*—describing events that happen completely over a brief period of time, i.e., what may be considered prototypical events). It was observed that in early L1 acquisition, English simple past tense marking (and its analogues in other languages) typically appeared with achievement verbs, or accomplishments verbs and rarely with activity verbs or stative verbs. That is, utterances such as *I drew it* or *I broke it* would be expectable at a relatively early stage of development, while *I played* and *I knew that* would not be expected at such an early stage.

This correspondence between past tense marking and inherent lexical aspect led to a hypothesis that has variously been called the *defective tense hypothesis* (Weist et al., 1984), the *primacy of aspect hypothesis* (Robison, 1990) and the *aspect hypothesis*

(Bardovi-Harlig, 1999). The original L1-related proposal was that early "tense" marking does not mark tense at all. Rather it redundantly marks verbs for their inherent aspect as accomplishments/achievements (sometimes combined under the label *telic verbs*). The temporal semantics allegedly involved in true tense inflection was held by Antinucci and Miller (1976) to be cognitively too advanced for L1 learners at the stage in question. However, their proposal was undermined by the demonstration by Weist and colleagues (Weist et al., 1984) that Polish children at the age in question were quite capable of cognitively representing events as being displaced in time and that they did in fact use activity verbs marked for past tense. In the debate that ensued (see Weist, 1986 and Andersen & Shirai, 1996) it was generally maintained that in both L1 and L2 acquisition there is a definite relationship between inherent verbal aspect (specifically, telic aspect) and inflections such as the English simple past, that is, inflections which mark events in adult native speaker narratives.

The other factor that has been argued to relate to the distribution of simple past tense forms (and the like) is discourse grounding. In relation to a mental model of a narrative, the foreground utterances are those which relate to the sequence of events that occur in the evolving model. A typical event involves a focal participant acting in such a way that the situation following the event clearly differs from the situation preceding it. Under the rubric of *transitivity*, Hopper and Thompson (1980) survey a collection of linguistic properties which tend to occur together, being united, they suggested, by their relationship to foregrounding. The ideal foregrounded event in their terms would be an event that in a brief span of moments is carried to completion by an active individuated subject behaving volitionally so as to strongly affect an individuated inactive object. These features of events in narrative have formal grammatical reflexes involving case, number, aspect, modality, etc. Among other things, this means that foreground sentences will be rich in telic verbs marked for simple past tense, perfective, or past perfective (depending on the particular language).

Flashner (1989) noted that in the English oral narratives of three L1 Russian learners of English, background verbs tended to be in their base-forms (i.e., simple, uninflected forms in the case of English), while foreground verbs were marked with the simple past tense inflection. She took this to be a transfer effect, with the English simple past inflection being used in contexts that would call for the Russian past-perfective, and the base-form being used in contexts that would call for the Russian past-imperfective. However, Bardovi-Harlig (1995) found evidence of the same pattern in the English spoken output of learners from L1 backgrounds other than Russian. In a cross-sectional study of thirty-seven L2 English learners, examining oral and written narratives, she found a clear tendency toward greater use of simple past tense marking in foreground sentences in contrast with background sentences. As the overall tendency to use past-marked forms (as opposed to base-forms) increased, the uneven distribution of past-marked forms continued to be the case. For example, learners who were providing the past tense marking in 40% of semantically past contexts were more than twice as likely to mark foreground verbs than background verbs, while learners who were providing tense marking in 80% of semantically past contexts marked past tense in 90.5% of foreground verbs as opposed to 72.5% of background verbs. Veronique (1987) found an analogous pattern of development among North African learners of L2 French.

Bardovi-Harlig (1994) points out the difficulty of distinguishing the factor of inherent verbal aspect from the factor of discourse background/foreground. Telic verbs are more typical foreground verbs than are atelic verbs. However, if inflectional morphology is understood as part of a cue system, the issue takes on a somewhat different appearance. As a grammatical cue, past tense inflection is learned when it is associated with some recurring variety of process involved in the construction or

modification of mental models. The past tense inflection thus first appears primarily with verbs which mark the completed steps of the narrative. This suggests that at the level of conceptual models of narratives, being a completed step is a salient characteristic. The story moves forward frame by frame. After a step in a narrative, things are different than they were before that step in the narrative. In Glenberg, Meyer and Lindem's (1987) mininarratives (Chapter 1, above), in the version in which the jogger took off the sweatshirt, the model contained a barebacked jogger after that step in the narrative. In the version where the jogger put on the sweatshirt, the model contains a sweatshirt-clad jogger after that step in the narrative. On the other hand, a background clause such as *while he was putting on his sweatshirt*, does not move the model to a new state of affairs, but rather creates an expectation of a separate step forward that will accomplish that, as in *While he was putting on his sweatshirt it tore*. This property of being a step forward in a story, a move from frame n to frame n+1, is apparently a conceptual property that can become associated with grammatical form.

Our assumptions thus point us to predict a sort of *primacy of aspect* regarding the distribution of past tense forms, but the aspect involved is not inherent lexical aspect but rather grammatical (discourse) aspect. That is, at the stage in which "past tense" morphology is associated with discourse foreground (that is, with steps forward in mental models), the so-called tense morphology is marking perfective aspect, not past tense. Virtually by definition of *foreground*, this will involve a high percentage of telic verbs. Since the use of the past inflection with foreground verbs arises after a period in which possibly no inflection is used, or else inflected forms are used in an unprincipled way (Housen, 1994; Andersen, 1991; Bardovi-Harlig, 1998), the systematic use of past (etc.) forms indicates that learning has taken place, that is that something has associated with something. Anderson has continued to maintain that part of what is involved is an association of the inflected form with the inherent lexical aspect of the verbs. That is, telic verbs, by assumption already categorized as telic verbs, come to be formally and redundantly marked as telic verbs by the past inflection. Andersen relates this to a proposed *principle of congruity*. From the vantage point of this dissertation, such a notion of redundancy needs to be viewed with caution. The conceptual feature *step in a narrative model* (that is, the incrementation of the model forward to a new situation) may be marked or not marked, but if it is marked, it is not redundantly marked. And truly redundant cues, that is cues that must coincide compatibly with other cues involved in the formation or incrementation of a mental model, (e.g., the plural marker in the phrase *three dogs*) are nevertheless fully active processing triggers, at least for native speakers. That is, the language processor expects these separate cues to converge, rather than simply ignoring, say, the "redundant" plural marker.

An important assumption made by researchers such as Anderson and Robison is that in applying English "past tense" marking (Spanish preterite marking, etc.) to telic verbs (and not to atelic verbs), the learners are not yet using the marking in a target-like (adult native speaker-like) way. Rather, only when the marking has spread to consistent use in "past contexts" with all verbs can it be considered a true tense marker. From our standpoint, the early use of an inflection with a telic verb as a marker of perfective aspect would appear to be a permanent development. If the conceptual property of being a step forward in a narrative model is salient enough to become associated with the co-occurring formal event of "past" inflection, how would the conceptual feature and the formal feature later become disassociated, given that they will continue to typically co-occur? Associations form as a result of co-occurrence in working memory. This particular aspect-like association of simple past tense with events in discourse will continue to occur consistently in an ever growing range of contexts. If the inflectional form marks perfective aspect early in the game, it will continue to do so. New, nonaspectual functions

might later be added to the old aspect-marking forms, but the old functions need not disappear.

2.5. STUDIES OF INFLECTION IN L2 RUSSIAN

Recent dissertations by Boots-Ebenfield (1995) and Nelson (1998) have addressed L2 Russian inflectional morphology, the former centrally, and the latter peripherally. Boots-Ebenfield examined transcripts of Oral Proficiency Interviews (OPIs) of thirty L2 Russian learners at the Intermediate and Advanced ACTFL levels (see Liskin-Gasparro, 1987 for discussion of OPIs and the ACTFL proficiency levels). For each learner there were two transcripts, one of an OPI conducted prior to an extended stay in Russia, and one of an OPI conducted afterward. Boots-Ebenfield tabulated all of the verbs in the transcripts, and examined aspectual form in relation to the proficiency level of the speaker, in relation to whether the OPI occurred before or after the stay in Russia, in relation to the aspect-related morphological category and aspectual value of each verb, and in relation to the function of each verb in discourse. He noted that with increases in proficiency, there were changes in the relative proportions of verbs belonging to different morphological categories (defined in terms of the morphological relationship between the perfective and imperfective forms of the same verb). Early on, there was a larger proportion of simple imperfective forms. Subsequently, there was an increase in the proportion of prefixed perfective forms. This increase occurred at the ACTFL Intermediate-High and Advanced levels, i.e., the levels at which learners increasingly produce simple narratives. In principle, the increase in the number perfective forms is in keeping with the fact of emerging narrative ability, since perfective is the aspect of narrative events. The final development observed in connection with morphological classes of verbs was an increase in the number of prefixed perfective verbs of the type that have (in the native lexicon) corresponding imperfectives derived by suffixation (so-called secondary imperfectivization). However, little use was observed of those corresponding derived imperfectives.

Boots-Ebenfield acknowledges that to some extent at least, the pattern observed could simply be the result of vocabulary growth. Basically, the rank order of type frequencies of the various verb categories in the corpora from the more advanced learners is similar to the rank order of frequencies of the verb categories that is found in the native lexicon. What appear to be largely lacking from Boots-Ebenfield's data are instances of the same verb being used in both aspects, varying in a nativelike manner according to discourse function. Therefore, an important question to ask is whether the changing pattern that was observed reflected increasing use of the aspectual system by the learners. It could be that the learners knew many verbs in only one aspectual form, with that form not fulfilling any aspectual function. For example, the word *vyras*, 'grew up' (grew.up:past:perfective:masculine) occurred in a nativelike perfective context, but it also occurred in a context where native speakers would use the imperfective form *vyrastal*. (From the data as presented, it is not even clear whether the different instances of *vyras*, in appropriate and inappropriate environments, are from the same or different speakers.) We can imagine a male speaker (the form is masculine) learning *vyras* in the context of giving biographical information about himself ("I grew up in Chicago"). Later, when needing to talk of people growing up in general, or of someone in the process of growing up, he might simply extend the perfective form to the new environment (an imperfective environment in native Russian). In the case of another verb, the perfective and imperfective forms do both occur in the corpus (*chitat'—prochitat'*, 'read') However, at least once, the imperfective form is also used in a perfective context.

In general then, it appears that Boots-Ebenfield's data are consistent with the following possibility: Learners initially acquire a single aspectual form of each verb. Although that form has aspect marking for native speakers, it fulfills no aspectual function for learners at that point. The lexical semantics (and closely related to that, the most characteristic discourse functions) of that verb determine the form in which each verb is most likely to be learned. Therefore, the form will naturally tend to be used in appropriate target language contexts. When the other aspectual form a given verb is needed for some context, it won't be available. That is, aspectual form is not being used contrastively. In those cases where there may appear to be contextually appropriate variation in aspectual form, it is possible that what is observed is formulaic speech. That is, the L2 speaker may have learned to use a particular form in connection with rather specific message content.

In other words, Boots-Ebenfield's data appear to tell us little regarding whether the Russian learners are acquiring the functions of those aspectual forms that appear in their speech. Now if in a spontaneous narrative, being related for the first time by a particular L2 speaker, the narrator said something such as *He fell off the cliff; while he was falling...* , in the process using the perfective verb form for the first instance of *fell,* and the imperfective form for the second (*falling*), that would raise the possibility that the aspectual contrast had begun to function for that speaker. Boots-Ebenfield's findings give us no such assurances. The research reported in the next chapter of this dissertation will help us to decide whether learners at the proficiency levels of those involved in Boots-Ebenfield's study might have been using aspectual forms for aspectual functions, or merely using the particular verb word-forms that they happened to know.

Nelson (1998) gathered extensive data from two Russian learners who progressed in proficiency from the ACTFL Novice-Low level to the Intermediate level through participation in an eight-week intensive course. Nelson argued in favour of a strong contribution of formulaic language to the spoken production of the learners (accounting for twenty to thirty percent of their utterances). It should be clear that mere production of inflected forms in early L2 Russian in no way demonstrates that inflectional processing is going on. If the learner knows a noun or verb for purposes of spoken production, s/he of necessity knows it in some inflected form, given that nouns and verbs are obligatorily inflected (even if with meaningful zero-marking). Thus, there are many inflected forms in Nelson's data. The inflection may have been non-functional, buried in formulaic speech. (Nelson points out that even a single word can be used in a formulaic manner.) In addition to the possibility of formulaic production of inflected forms for particular frequently repeated message-level units, there is the possibility of learners producing some inflected forms by means of conscious metalinguistic planning.

Both Boots-Ebenfield and Nelson leave us in the predicament of seeing in L2 spoken production lots of what, from the native perspective, would be inflectional morphology. However, from the perspective of the learners' developing language systems, the particular formal features may or may not be actually functioning. The experiments to which we turn in Chapter 3 will hopefully give us a better idea of the level of inflectional processing that may commonly be possible for learners at the proficiency levels of those in Boots-Ebenfield's and Nelson's studies.

In this chapter and the previous one, we have created a framework of understanding in which to examine the development of sensitivity to inflectional morphology in a second language. In the following chapter we proceed to do just that.

3. Sensitivity to Inflection in L2 Russian

In Chapter 2, we noted that the ability to make use of inflectional morphology can be compromised in aphasia (Luria 1946, cited by Alpatov, 1997; Smith & Bates, 1987; Frazier & Friederici, 1991). The patients described by Luria processed sentences as though case marking were not present, identifying a word form as a member of a particular lexeme in order to understand the sentence, relying on semantic properties of lexemes, but remaining insensitive to case-marking that conflicted with the semantic cues. The case cue appeared to be inert for such patients, failing to trigger processing related to the assignment of semantic roles or discourse roles.

As children learn Russian, their acquisition of case marking accompanies growth in vocabulary, increasingly adult-like pronunciation, increasingly long and complex utterances, and so forth. For example, Gvozdev (1949, cited by Slobin 1966b) recorded the child utterance shown in (1), where (2) is the adult analogue (example provided by Dan Slobin, p.c.):

(1) glaj-u mamatsk-om
play-1pers mamma-instr
I am playing mama.

(2) Igraj-u s mamochk-oj
play-1pers with mama-instr
I am playing with mama.

The child, just over two years of age, shows productive use of the first declension instrumental ending, using it with a second declension noun (which employs the ending *–oj* rather than *–om* in adult speech). The child also omits the preposition *c*, 'with', omits the initial syllable of the first word and mispronounces at least two segments. In other words, case marking seems to be taking hold at an early stage, long before the child's speech is very adult-like. This implies, of course that the instrumental case has been "understood" by the child. That is, when nouns ended with *–om* in this child's experience, they often had the sociative role (action carried out together with another actor) and this role became mentally associated with that ending, so that the ending now shows up in speech on nouns whose referents have the sociative role at the conceptual level. That ending will continue throughout life to frequently trigger that role-assignment during comprehension (unless the system becomes physically damaged). The production system will learn to automatically supply the instrumental form for message level sociative participants. The use of instrumental case as a comprehension cue and its suppliance in speech, may become partly independent of one another. Thus Smith and Bates (1987) present evidence that there can be some loss of the ability to use inflectional form for comprehension without corresponding agrammatic speech production, and also cite patients for whom the opposite appears to be the case—agrammatic spoken production without detected problems in the use of grammatical morphology in comprehension. This is compatible with the idea presented in Chapter 1 that the comprehension system trains the production system (Stemberger, 1998). Once trained, the production system may be able to go on functioning without the support of the comprehension system. On the other hand, Smith and Bates (1987) do find that in general there is an direct relationship between the severity of impairment in inflectional processing during comprehension and the presence of corresponding agrammatic spoken

production, possibly suggesting at least some ongoing relationship between comprehension ability and production ability.

In any case, it is difficult to imagine an inflectional form emerging in speech with a function approximating the adult function before that form has begun to function in comprehension. Turning to adult second language learners, one would hope to see evidence that aspects of inflectional morphology are becoming active (as opposed to remaining inert) during comprehension. Highly active, nativelike use of inflection for comprehension should cause the learner's own nonnativelike productions to clash with the same learner's increasingly nativelike comprehension system, thus allowing the production system to learn to be increasingly nativelike, or at least making the learner aware of nonnative features in his or her own spoken output.

When the child cited above produced the instrumental case where it was appropriate, and particularly given the fact that the first declension form was overgeneralized to a second declension noun (ruling out the possibility that that the child was producing the form within a formulaic utterance), it clearly meant that the child's comprehension system had been dealing with the instrumental case as well. The adult language learner, unlike the child, might be able to mimic such features of native spoken output through metalinguistic planning of utterances. Thus observing an adult L2 Russian learner using the instrumental case in an appropriate context does not tell us whether that learner's comprehension system has come to "understand" the instrumental case in such contexts. That is, we do not know whether the instrumental case triggers and/or constrains comprehension processes appropriately when encountered in such contexts in the flow of speech. In order to know whether inflectional forms are active or inert we need a way to tap into their operation in the course of comprehension.

The approach taken here is similar in spirit to Luria's presentation of bizarre sentences discussed in the previous chapter. He created sentences in which case cues were deliberately made to clash with other types of cues (lexical cues, semantic cues, word order cues). If L2 users were to experience such a clash during meaning-focused listening, this would provide evidence that some of the other types of cues were used *and* that the case cues were also employed. In order for cues to clash, they must be active. The same basic logic can be applied to other types of inflection as well. Thus the aim in the experiments reported in this chapter was to elicit responses that reflect the experience of a processing clash on the part of the listener, rather than simply reflecting the application of metalinguistic knowledge in an analytical fashion.

Kempe and MacWhinney (1998) attempted to do this by having participants in their study listen to noun pairs embodying combinations of cues, including word order, semantics and case-marking, after which the participants would select a picture that corresponded to their understanding of the noun pair (with a constant verb being understood). The pictures differed in terms of which participant was depicted as agent and which was depicted as patient. This was intended to enable the experimenters to examine the extent to which case-marking was used in comprehension. This approach may not, however, prevent the adoption of metalinguistic strategies. For example, a participant (having already been trained on the nominative forms of the nouns in a pair) could choose as object any noun that differed from the nominative form. It would be helpful to know whether nonnative Russian users at the proficiency level of the participants in Kempe and MacWhinney's study in fact make use of case cues during normal comprehension, and to what extent they make use of them, before agreeing that the participants were employing case inflections as normal comprehension cues in the experimental task.

There is much to be said in favour of using an implicit measure of sensitivity to inflectional errors such as the word-monitoring paradigm used by Friederici et al. (1992)

in the investigation of inflectional processing by agrammatic patients. Under the constraints of the research environment, this did not prove practical for the present study. Instead, an approach was taken in the spirit of Luria's simpler demonstration of insensitivity to inflectional morphology: present L2 users with sentences that are relatively easy to process semantically; require an immediate reaction to the meaning of the sentences; limit opportunities for metalinguistic reflection; and include occasional sentences in which an inflectional cue clashes with other cues. If an L2 listener does not notice the clash, then we can conclude that the inflectional cue was inert (insofar as it was perceived). If the listener does react to the clash, it is not certain proof that the form was processed linguistically. There is still the possibility of metalinguistic decision-making, but this possibility is greatly reduced. It seems out of the question that metalinguistic knowledge could regularly be employed as a means of making inflectional cues active during the rapid flow of speech, even by those who are able to use such knowledge for planning certain features of their spoken production. We can support this intuition empirically by including some measure of metalinguistic knowledge and demonstrating that there is not a strong relationship between metalinguistic knowledge and on-line cue use. Therefore, in the experiments reported below, the sentences containing cue clashes were presented three times under conditions that increasingly allowed for the use of metalinguistic knowledge regarding grammatical form.

In connection with Russian inflectional morphology, if every possible variable were taken into account, there would be perhaps hundreds, if not thousands, of possible experimental comparisons. An attempt was made to sample a range of inflectional categories from the vantage point of the conceptual side of inflectional processing. There was a desire to encourage language processing that would be comparatively normal within the constraints of such an experiment. Thus Tamara Alexandrovna Ivanova constructed connected, sensible (if simple-minded) texts first, and then we explored them for contexts in which to introduce errors, making only minimal changes. In the two experiments reported below, an effort was made to test for sensitivity to the following types of errors:

(A) An oblique case substituted for the nominative case
(B) The nominative case substituted for an oblique case
(C) One oblique case substituted for another oblique case
(D) The nominative case substituted for the accusative case
(F) The accusative case substituted for the nominative case
(G) Perfective aspect substituted for imperfective aspect
(H) Imperfective aspect substituted for perfective aspect
(I) Inappropriate verbal gender agreement
(J) Inappropriate verbal person agreement

The term *oblique* is used here in the sense of nouns with thematic roles that the verb of the clause is not subcategorized for, (i.e., which do not fill a syntactic argument position). This excludes direct and indirect objects, as well as goals of dynamic verbs in Russian. It was suspected that a strong contrast would be created by substituting incorrect nominative forms for correct oblique forms and vice versa. The oblique cases that were included were the instrumental and the locative cases (dative case not being fundamentally oblique and genitive case also having a somewhat special configurational status). This also provided the opportunity to contrast an oblique case that is governed by prepositions (the locative case) with one that carries the entire functional load of thematic role marking (the instrumental case).

This set of error types was chosen following piloting of an attempted experiment with a larger set and a larger number of exemplars of each type. That experiment proved

far too demanding for practical administration. The aim here was to create experiments that could readily be conducted both with individuals, and with groups of students in classrooms.

Two sorts of extraneous variables (among others) might affect the ability of nonnative speakers to react to errors: (1) syntactic or discourse contextual factors that increase overall processing demands; and (2) factors that increase or decrease the salience of the error. The expectation was that the effect of error type (understood in terms of the processing functions of the inflectional categories) would be powerful enough to be detected in spite of the presence of the second class of variables. In addition, the texts were designed to include relatively basic language, with visual pictorial support. Finally, in the placement of errors, the salient sentence initial and sentence final positions were avoided, as were inflectional changes involving stress alternations.

Two hypotheses with respect to case-marking were as follows: (1) The substitution of nominative case for oblique cases would cause fewer reactions than the substitution of oblique for nominative or oblique for oblique. (2) The nominative-accusative contrast would be more robust than other contrasts. The first expectation was in part based on the intuitions of native Russian teachers of Russian as a Second Language that learner errors in spoken output follow this principle. In particular, it was felt to be common for nominative forms to be substituted for non-nominative forms in learner speech and for non-nominative forms to be substituted for one another, but rare for oblique forms to be substituted for nominative forms. One explanation of this pattern in spoken production would be that it (at least in part) reflects the sensitivities of the developing L2 comprehension system.

The second expectation was based on the intuition that there is something especially fundamental in the subject-object distinctions, where both nouns are topical and less consistently characterizable in terms of semantic role than in the case of oblique nouns. This is related to multiple functions involved in the implementation of the case morphology as discussed in the previous chapter. Oblique cases are more purely semantic in nature than the nominative case. There is a well-defined inflectional system used for locative contexts. Furthermore, a noun marked for instrumental case, when not governed by a preposition and when adjoined to an active verb phrase, has a basic function of marking instruments.

In contrast with oblique cases, the nominative case (and to a lesser degree the accusative of direct object), does not have such a purely semantic-role-marking function. Rather, nominative case functions to mark the focal participant, which may in fact have a variety of semantic roles. Thus, it was reasoned, the nominative-oblique contrast used here (nominative versus locative, nominative versus simple instrumental) provides a contrast between fundamentally different linguistic functions of case marking. Likewise the nominative-accusative distinction, when used to distinguish subjects and objects, involves the unique function of distinguishing between focal participants (primary and secondary topics in Givón's terms). Detecting different developmental patterns in the presence of such strong functional differences might be easier than detecting differences in the developmental patterns of, say, two oblique cases.

The Prague School notion of markedness as further developed by Greenberg (1966; Croft 1990) is relevant here. In the Greenbergian sense, the oblique forms might be considered similar to one another in markedness (less frequent, more regular, more neutralization, etc.), while more marked than the nominative forms (which are the unmarked forms).[15] The nominative forms are also less marked than the accusative forms

[15] In Greenbergian terms, the unmarked character of the nominative case is reflected in the following properties: For the first and third declensions, the nominative form is the zero form, and with

(and accusative-oblique comparisons do not enter into this study). A plural example was included among the errors (actually a pluralium tantum—*br'uki*, trousers, incorrectly presented in a nominative case context in the locative case form *br'ukax*). In Greenbergian terms, plural forms are more marked than singulars, reflected in Russian by the fact that, among other things, there is a substantial reduction in morphological variety in the plural oblique forms when compared with the singular forms. However, in the case of this specific error, there is the same *relative* markedness difference as in other errors. That is, there is a difference of a single value of an inflectional category (case). An especially important fact related to markedness is that crosslinguistically, marked forms tend to be longer (or otherwise more phonologically salient) than the unmarked forms. This could make it difficult to avoid confounding differences in acoustic salience with differences in case function.

In relation to verbal aspect, the major expectation was conditioned by experience with samples of L2 Russian spoken production. While conducting proficiency testing, I discovered a tendency of learners to overuse imperfective forms for foreground events in narratives at an earlier stage of development and an increase in the appropriate use of perfective forms at a later stage. It was therefore expected that participants in the experiments would be more likely to react to the substitution of incorrect perfective for correct imperfective than they would be to react to the opposite substitution. In view of hypotheses discussed earlier regarding the relationship of aspectual choice to lexical aspectual semantics and discourse grounding, we might attempt to derive and test other predictions. That is, we might predict that contextually inappropriate aspect involving a telic verb with imperfective aspect or a non-telic verb with perfective aspect would tend to trigger more reactions (under the primacy of aspect hypothesis discussed above) than would contextually inappropriate but nevertheless lexically congruent aspect marking. Alternatively, we might predict that inappropriate use of a perfective form with a background verb, or imperfective form with a foreground verb, would trigger reactions more reactions than contextually inappropriate aspect in which grounding and aspect were congruent. However, we noted that these factors (lexical aspectual semantics and discourse grounding) tend to be confounded in text, and it would have been difficult, if not premature, to use the experiments reported below in an effort to choose between these two hypotheses. At this stage of research, we are primarily interested in determining whether nonnative speakers are sensitive to aspectual errors (and therefore, are potentially processing aspectual inflection). Our desire to detect such sensitivity led us to create errors that might be expected to maximize the likelihood of reactions. Thus, our aspect errors involve violations of *both* proposed functions of aspect.

In a rather indirect way, our findings could turn out to bear on the choice between the primacy of aspect hypothesis and the discourse-based hypothesis: the discourse-based hypothesis would predict that sensitivity to aspectual errors would be a relatively late development, coming at a stage when the developing language processor is able to deal with extended, textured discourse. Lexical semantic properties, by contrast, are already present in single utterance discourse. Therefore, if the evidence should point to extremely delayed development of aspectual processing, that might be taken as support for the discourse-based hypothesis regarding the functions of aspect, though rather weak and indirect support. On the other hand, if it turns out that learners show either no sensitivity to aspectual errors, show strong sensitivity, or show sensitivity that appears to change over time, that will bear directly on the questions raised in this dissertation

---

the exception of the genitive plural, the nominative form is never longer than other forms; in neutralized, "nondeclined" contexts, the nominative case is used; the nominative case shows at least as much morphological variety as other cases; the nominative case is the most frequently occurring case.

regarding the development versus lack of development of inflectional processing mechanisms.

The issue of markedness in the case of Russian aspect is especially complicated and is best understood in terms of local markedness (Tiersma 1982). An unmarked imperfective lexeme can be marked as perfective by the addition of a prefix, while an unmarked (typically, prefixed) perfective lexeme can be marked as imperfective by the addition of a suffix (see the discussion in the previous chapter). In the experiments discussed below, the plan was to include one example of a perfective verb used in an obligatorily imperfective context and one of example an imperfective verb used in an obligatorily perfective context. In terms of inherent lexical aspect, the perfectivized imperfective involved a verb denoting a state, *stojat'*, 'be standing', while the imperfectivized perfective involved a verb denoting a brief event, *zakryt'*, 'close'. Morphologically, both errors involved the addition of an affix. The prefix *po* was added to the form *stojat'*, 'be standing' to create the perfective *postojat'*, 'stand for awhile' (called a delimitative form; see Durst-Anderson, 1992, p. 25). The suffix *-yva-* was added to the rood *zakr-* of *zakryt'*, 'close' to form the so-called secondary imperfective *zakryvat'*, 'be in the act of opening/open periodically (etc.)'. In fact, it proved difficult in the earlier abandoned experiment and in the present experiment, to substitute one aspect for the other in such a way as to produce an error, rather than a mere change in grounding (e.g., from foreground to background, or vice versa). There was some doubt regarding the status of aspect substitutions used in the present experiments. Because of this, the imperfective-for-perfective substitution was presented twice in the first experiment. In the case of the perfective for imperfective substitution (*postojat'* for *stojat'*, 'be standing'), my primary consultant and the creator of the texts used in the experiment, Tamara Alexandrovna Ivanova, considered the substituted word-form to involve a lexical derivation and not a pure aspectual change. Russian dictionaries were not consistent on this point. The substitution certainly involves an aspectual change in the general, crosslinguistic sense of the term, and in any case, the resulting verb is perfective. In context, this appears to create an impossible meaning when comprehended in a nativelike way (for one informant connoting a temporary cessation of walking on the part of the furniture mentioned in the text). Not surprisingly, native controls reacted strongly to this error (evidenced by facial expressions of amusement). They also reacted uniformly to the two imperfective-for-perfective substitutions in the first experiment.

Specific predictions:

In the introduction three questions were raised. We are now in a position to more meaningfully state specific hypotheses which address these questions.

Question 1: Do adult learners of L2 Russian develop the ability to make use of inflectional morphology in on-line comprehension?

A finding of a clear effect of *group* (Intermediate vs. Advanced) in the experiments reported in this chapter (especially the first task of both experiments) would provide support for an affirmative answer to this question. In particular, such a finding could be viewed as a necessary (though not sufficient) condition for claiming the existence of inflectional learning. Of course, a significant difference between the groups would not give any assurance that the learning will ever go on to completion, but rather would give hope that it is at least under way. In testing the null hypothesis (no difference between groups) therefore, we are attempting to support the following claim.

Hypothesis 1: L2 Russian users with greater experience in the language will show more sensitivity to inflectional cue-clashes than those with less experience.

There are various reasons why such an outcome might not be expected. Gass (1997) follows Swain (1985) in suggesting that whereas spoken production forces syntactic processing, listening comprehension can be relatively successful while ignoring syntax and relying on semantic processing. This certainly raises the question of whether syntactic processing would ever develop in listening comprehension. In connection with the adverb-placement issue with French L1 learners of L2 English, White (1989) divided her participants into two proficiency levels, and found no effect of proficiency on performance related to that notoriously fossilizable aspect of L2 English. Thus, it is not unreasonable, especially in view of the complexity of Russian inflection, to ask whether there are detectable differences between less advanced and more advanced groups as regards their sensitivity to inflectional cue clashes (and by inference, their tendency to process inflection) under conditions where the application of metalinguistic, analytical knowledge is impossible. We certainly cannot assume that frequent suppliance of inflectional form in spoken production is an indication of native-like functioning of inflectional form in comprehension (or in production, for that matter). Rather, comprehension processes need to be investigated in their own right.

Evidence for increasing sensitivity to inflectional form over time would not yet tell us anything about the nature of the developments. It is important to investigate whether any developmental pattern might be detectable. Systematic variation in the levels of sensitivity to various inflectional categories, if it were to follow linguistically characterizable principles, would encourage the hope that developmental differences between groups not only reflect learning, but reflect *language development* in an interesting sense. Thus, the following question would naturally follow from findings in support of Hypothesis 1.

Question 2: Is there a pattern in the development of sensitivity to inflection, and if so, what is its nature?

Given a null hypothesis of random patterning of sensitivity to errors, we might predict the alternative hypothesis:

Hypothesis 2: There will be systematic patterning in the tendency of participants to notice some errors and error types earlier and more often than others.

In addition, this tendency should be characterizable in terms of functional considerations of the sort considered in Chapters 1 and 2. That is, although the hypothesis address the issue of whether there is systematic patterning in the development of inflectional processing in L2 Russian, we will also be highly interested in what that patterning might look like. That aspect of the research is carried out more in an exploration mode than a hypothesis testing mode.

The third question raised in the Introduction was the following:

Question 3: What, if any, is the relationship between metalinguistic (analytical, formal grammatical) strategies and the sensitivity to inflection that is evidenced in on-line listening?

The outcome of the experiments might bear on this question in the following way. For starters, the first listening task is intended to eliminate any possibility whatsoever of metalinguistic analysis-based decisions, so that reactions to cue clashes can be assumed to reflect language processing mechanisms. Then the second and third tasks are intended to allow increasing scope for the application of metalinguistic analytical strategies. Therefore an effect of *task* would be consistent with an expectation that the ability to perform metalinguistic analytical judgments regarding linguistic form taps different processes from those involved in on-line listening comprehension. This leads to the following hypothesis:

Hypothesis 3: Error detections will increase with increasing opportunity to apply metalinguistic analytical knowledge.

An effect of task could have other explanations, and so here again, findings in support of Hypothesis 3 will fall into the category of necessary but not sufficient findings.

It is reasonable to expect that many participants would perform strongly in a pencil and paper task involving the identification and correction of (pedagogically) elementary inflectional errors. All participants had received formal grammar instruction at some point during their Russian learning. However, it is not a foregone conclusion that performance in an unconstrained pencil and paper task will be superior to performance in the listening tasks. In fact, Tarone (1985) found an analogous prediction to be false with respect to certain grammatical elements of L2 English in spoken production. That is spoken output appeared to be more nativelike than did performance in a pencil and paper task.

The comparison of the listening tasks with a less constrained pencil-and-paper task could also be of interest in connection with the question regarding whether metalinguistic judgment tasks (of which the pencil and paper task will be a clear example) rely on linguistic mechanisms or rely on more general cognitive mechanisms (Masny & d'Anglejan, 1985; Birdsong, 1989). That is because we will be able to see the extent to which performance in an unconstrained metalinguistic task exceeds performance in tasks where metalinguistic reflection and problem-solving are less possible or impossible.

The question of the possible interdependence vs. independence of metalinguistic analytical knowledge and listening comprehension mechanisms is not addressed by Hypothesis 3. However, this is another important matter for exploration, and therefore the correlations between the pencil and paper tasks and the listening tasks will be examined, in an effort to shed further possible light on findings related to question three.

3.1. EXPERIMENT 1

3.1.1. Method

3.1.1.1. Participants

Thirty-seven nonnative speakers of Russian residing in St. Petersburg, Russia, participated in the experiment, as well as thirteen native Russian control participants. The grouping criterion was not determined until after the results from the first sixteen experimental participants had been examined, and that criterion requires some explanation.

The initial sixteen subjects divided naturally into those who had been learning Russian approximately two years and those who had been learning Russian four years or more. The variable *time since onset of Russian learning* appeared to be a more promising a predictor than the variable *time of residence in Russia*. Based on that early evidence, the criterion chosen for assignment to groups was *time since the onset of Russian learning*, with four years constituting the dividing line between groups. This criterion was then applied rigidly to the remainder of participants in Experiment 1, as well as all of the participants in Experiment 2. It seems extremely likely that by more careful criteria some participants would have been grouped differently than they were. It should be kept in mind that the group label *Advanced* could be more literally stated as *≥ 4 years' learning*, while *Intermediate* could be more literally stated as *< 4 years' learning*. The notion of a pseudolongitudinal design as discussed by Gass and Selinker (1994) involves assignment to groups based on some measure of proficiency. In such a design, differences in measured proficiency are taken as a reflection of different points along an idealized learner time-line. Here the relationship of the groupings to the temporal axis is actually more direct. Obviously, such a strategy would not be advisable in every situation. In the present case, all participants were functional, regular Russian users, and viewed themselves as long term learners.[16]

The mean length of time reported by the *Advanced* group was 5.79 years, ranging from four to fifteen years. Sixteen of the twenty in the Advanced group had been learning Russian from four to six years. Keep in mind that the participants in both groups were not only Russian learners, but also serious users of Russian in a second language context. In the case of the Intermediate group the mean time reported since onset of learning was 2.15 years. Only one participant had been learning Russian less than a year, and two reported 3.5 years. The average time in Russia for the Intermediate group was 1.65 years (only one participant having lived in Russia less than six months) and for the Advanced group 3.6 years.

The L1s of the participants included English: 17 (9 Intermediate, 8 Advanced); Chinese: 8 (3 Intermediate, 5 Advanced), Finnish: 3 (2 Intermediate, 1 Advanced), Spanish: 2 (Advanced), German: 2 (1 Intermediate, 1 Advanced), Farsi: 1 (Advanced), Italian: 1 (Advanced), Japanese: 1 (Intermediate), Korean: 1 (Intermediate) and Turkish: 1 (Intermediate). There were seventeen males and twenty females. Ages ranged from mid-twenties to early fifties. A few participants had begun Russian study in a high-school, but for most or all it can be assumed that serious Russian learning began in late adolescence or adulthood.

A large proportion of the participants were students, including undergraduates (23), graduate students (4) and students in programs of Russian for foreigners (7). The others were people who regularly use Russian in their everyday lives, some quite extensively, and all of whom have studied basic Russian in formal courses at some time early in their learning of Russian.

3.1.1.2. Materials

Two texts were constructed to accompany picture stories taken from Takahashi & Frauman-Prickel (1985). One depicted a man walking to a bathroom sink and washing and drying his hands (ten pictures in all). The other depicted the same man watching a television, turning it off, going to bed and falling asleep (eighteen pictures in all, one excluded, leaving seventeen). The texts were created for this experiment by Tamara

[16] The validity of the grouping is ultimately to be supported, or cast in doubt, based on whether it leads to a rejection of the null hypothesis in connection with Question 1.

Alexandrovna Ivanova of the St. Petersburg State University. They were read aloud and tape-recorded by another female adult native speaker. Both texts are presented in printed form in the Appendix. The texts were broken up into segments corresponding to the pictures. Each segment contained one or more sentences. With both texts the first segment was particularly long, as it set the stage for the rest of the discourse.

Before they were tape-recorded, the texts were modified by the introduction of inflectional errors (indicated in the Appendix by boldface, followed by the correct target form in parentheses). An error was introduced into the first segment of the first text for training purposes and this item was used to train participants in the experimental task. Counting the one in the training segment, there were a total of ten errors in twenty-seven sentences, eight of which were tabulated, with not more than one error per text segment. In no case was an erroneous word the initial or final word of a sentence. The distribution of errors was not deliberately randomized, since it depended in part of the possibilities provided by the context. However, the distribution of errors and error types has every appearance of being unsystematic.

For the first task of the experiment (error detection during meaning-focused listening), photocopied pictures from Takahashi & Frauman-Prickle (1985) were provided as the basis of a concurrent multiple choice picture selection task. For each text segment, three pictures were provided in a row, the pictures being designated as a, b and c. Only one of the three pictures in the row was appropriate to the meaning of the corresponding text segment. Thus there were twenty-seven numbered rows of pictures for the twenty-seven text segments, with three pictures, a, b and c, in each row. One of the pictures chosen as a distracter in each set of three was clearly wrong. The second distracter picture contained overlapped with the correct picture in some component or components. The pictures were arranged so that turning pages would not interfere with concentration on any item that contained an error. The purpose of this concurrent task was two-fold. First, it was intended to insure that participants fully processed the texts for meaning. Second, the need to search for and identify the correct picture added a divided attention feature to the task, intended to increase the likelihood that reactions to errors would be the result of unconscious, automatic processing mechanisms. The concurrent task was not intended as a measure of comprehension. However, an overall low performance in this task would indicate that the attempt to construct adequately simple texts had been unsuccessful.

Answer sheets with instructions were provided for each task (see the following section). For the third task, the answer sheet contained, in printed form, just those sentences from the texts that contained errors. This task was intended to be a simple one for any participant who possessed the relevant formal grammatical (metalinguistic, analytical) understanding of the inflectional categories in question.

3.1.1.3. Procedure

In all three tasks, there were written instructions, which were supplemented by an oral paraphrase, and oral clarifications were provided in response to questions that were raised by participants. The need to refrain from guessing and only indicate errors when absolutely certain was always strongly emphasized.

The experiment was conducted individually with three participants and otherwise in groups numbering from three to eight individuals. An effort was made to keep the listening conditions as similar as possible. In general the entire experiment took from twenty to twenty-five minutes, including explanations and handing out and collecting materials, with larger groups taking longer than smaller groups for organizational reasons.

3.1.1.3.1. Task 1: Meaning-Oriented Listening

In this task, participants listened to the tape recording of the text segments. For each segment they indicated, by writing a, b, or c, which picture in the row of three pictures was the one being described in the text. The first segment was used to illustrate the task and briefly discussed. The participants then listened to the remaining text-segments, each time indicating in multiple choice format which picture corresponded to the segment they heard. In addition, they were instructed to indicate by an "X" any segments in which they heard an obvious mistake. During this task, there was a six second pause after each text segment.

3.1.1.3.2. Task 2: Form- Oriented Listening

In this task, participants listened once again to the same twenty-seven text segments. This time there was no picture selection task and the pause after each segment was lengthened to ten seconds to allow silent reflection on the form of the segment. Participants indicated with an "X" that a particular segment contained an error.

3.1.1.3.3. Task 3: Printed Sentences

For this task participants were provided with just those segments which contained errors. They were instructed that each segment contained one and only one error and that they were to find the error and correct it. This was intended to allow the greatest possible opportunity to bring explicit metalinguistic knowledge to bear on finding the errors. No specific time limit was placed on this task.

3.1.2. Results

The thirteen native Russian speaking controls performed only the Meaning-Oriented task. As a group, they noticed 98% of the errors. Only one error was missed, by two participants, one of whom was observably distracted at the relevant moment. Eleven of the thirteen participants thus detected 100% of the errors. Two items received unexpected error judgments (segments 8 and 13 of Text 2 each received error judgments from four of the native controls). One control participant, who subsequently reported equivocating over these, stated that the problem was lexical rather than grammatical. In any case, these two items were excluded from the tally of false alarms (incorrect guesses). The mean score of the native controls in the picture selection task was 99.7%.

Turning to the experimental participants, there was one case when a participant reacted out loud to an error in a group of seven participants (out of whom six then marked the error). The error in that item (segment 7 of Text 1) involved the substitution of an imperfective for a perfective verb form. Fortunately, it turned out that this was the error that was repeated in the experiment (nineteen segments later). Therefore, segment 7 of Text 1 was removed from the analysis of the experiment. Segment 16 of Text 2 was retained, but may have received an inflated number of number of responses as a result of this incident involving the identical error in segment 7 of Text 1.

In the Meaning-Oriented and Form-Oriented tasks, indications of errors were only counted when the Printed Sentences task demonstrated that the participant was able to identify the precise error involved. This reduces the number of possible cases where participants might have marked an "X" correctly in the two listening tasks as a result of guessing. This affected only 2.9% of the correct responses in the Meaning-Oriented task, and 4.5% in the Form-Oriented task. In other words, the proportion of items where participants correctly claimed there was an error, but were subsequently unable to identify

the error and make a credible (and in all but a very few cases accurate) attempt at correcting it, was marginal.

Looking at the two groups and three tasks as a whole, an analysis of variance revealed significant main effects of both task ($F[2,35] = 119.521$, $MSe = 1.874$, $p < .001$) and level (Advanced vs. Intermediate), ($F[1,35] = 10.981$, $MSe = 5.856$ $p < .01$) and no interaction of task by level ($F[2,70] = 1.013$, $MSe = 1.874$, $p = .368$) as summarized in Table 1.

Source	df	Sum of Squares	Mean Square	F-Value	P-Value
Group	1	72.489	72.489	12.379	.0012
Subject	35	204.953	5.856		
Task	2	448.013	224.007	119.521	.0001
Task * Group	2	3.797	1.898	1.013	.3684
Task * Subject	70	131.194	1.874		

Dependent: Task

Table 1. Experiment 1: ANOVA source table, three experimental tasks performed by two groups (Intermediate and Advanced speakers of Russian, as defined in text), with participants as the repeated measure.

A separate analysis was performed in which experimental items rather than participants were used as the repeated measure (percentage of participants detecting each error was used as the dependent measure, rather than absolute scores due to the uneven group sizes). Again the effects of task and group were significant (for task, $F[2,21] = 65.440$, $MSe = 61.509$, $p < .001$) and for level ($F[1,21] = 80.330$, $MSe = 232.687$ $p < .001$). and the interaction of task by group was not significant, [$F[2,21] = 2.169$, $MSe = 61.509$, $p = .139$).

Figure 1 summarizes the findings graphically. We turn next to specific comparisons.

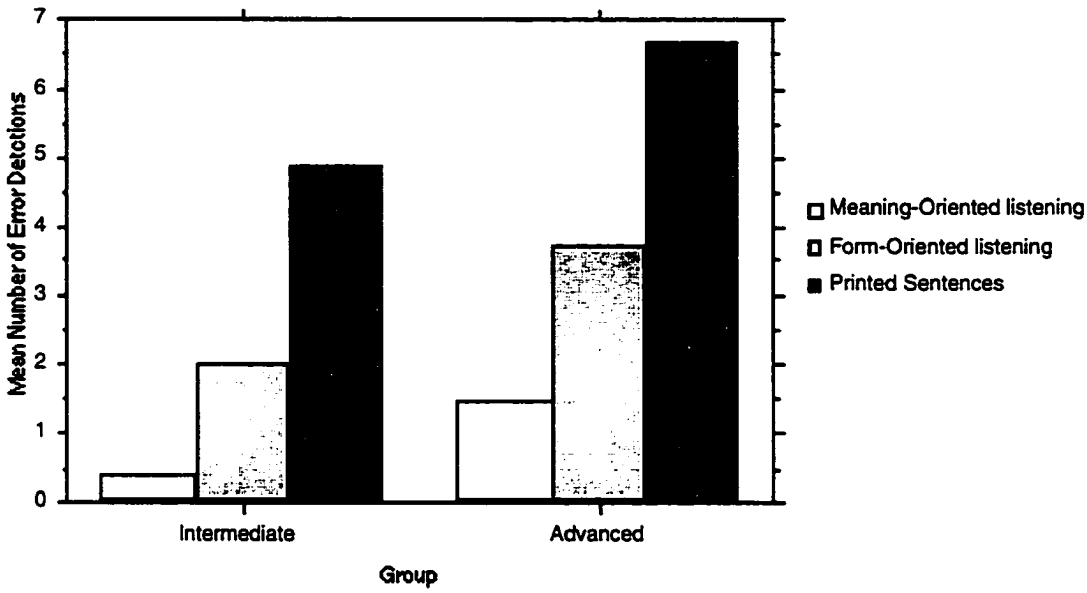

Figure 1. Experiment 1: Error detections in three tasks by the two groups. (The maximum possible number of error detections was eight.)

In the concurrent picture selection task, the mean score for the experimental participants was 91.7%. Recall that concurrent task was not intended as a measure of comprehension. Rather, an effort was made to construct texts that would be generally comprehensible to participants, since the purpose was to insure that participants' would fully process the meaning and would at the same time be prevented from engaging in metalinguistic reflection. The demanding pace of the task made it likely that there would be some errors in picture selection that were not due to an inability to comprehend. Nevertheless, the performance in this concurrent task was strong enough to assure us that the texts were in fact relatively comprehensible to most participants. Interestingly, the mean correct scores of the Intermediate group was nonsignificantly higher than that of the Advanced group (93.6% and 89.8% respectively). This further reinforces the conclusion that the texts represented a relatively low level of difficulty, as desired.

3.1.2.1. Meaning-Oriented Listening

In general, participants tended not to detect errors during Meaning-Oriented listening. The average number of errors detected overall was .946 (11.8%) out of the possible eight errors. The mean number of errors noticed per participant for the Intermediate group was .375 (4.7%) out of the possible eight (SD = .62, range from 0 to 2). The mean number errors noticed per participant for the Advanced group was 1.45 (18.1%; SD = 1.64, range from 0 to 5). Only four of the twenty Advanced participants detected fifty percent or more of the eight errors and all the others detected twenty-five percent or less. The difference between the Intermediate and Advanced groups on the Meaning-Oriented task was significant (t[35] = 2.61; p = .013, 2-tail).

3.1.2.2. Form-Oriented Listening

In the Form-Oriented task, participants detected an average of 3.027 errors (37.8%) out of the eight possible errors (*SD* = 2.179, range from 0 to 7). This represents an increase of 2.081 (26%). In the Form-Oriented task, the Intermediate group detected an average of 1.94 errors (24.2%) out of eight possible (*SD* = 1.18, range from 0 to 4), or a gain of 1.57 (19.6%) errors per participant over the Meaning-Oriented task. In the Advanced group the mean number of errors noticed was 3.9 errors (48.7%) per participant (*SD* = 2.47, range from 0 to 7), or a gain of 2.45 (30.6%) per participant over the Meaning-Oriented task. Although the interaction of task by group was not significant, nevertheless the Intermediate participants displayed a five-fold increase in the average number of errors detected, while the average number detected by the Advanced group doubled. In this sense, the increased time and decreased pressure (and the fact that this was the second hearing of the same text) were in fact more crucial to the Intermediate participants than to the Advanced participants. On the other hand, the absolute gain of the Advanced participants was greater than the absolute gain of the Intermediate participants, suggesting that perhaps the Advanced participants were able to make better use of that increased time and decreased pressure (and repetition) than were the Intermediate participants.

Seven of the seventeen Intermediate participants detected three or four of the eight errors in this task. Eleven of the twenty Advanced participants noticed four to seven of the errors, including five who detected seven of the eight possible. The difference between groups was significant ($t(35) = 2.90$; $p < .01$, 2-tail).

3.1.2.3. Incorrect error judgments (false alarms)

An important finding in the first two tasks is the incidence of false alarms, participants indicating errors when none were present. In the Meaning-Oriented task, there were an average of .973 false alarms per participant and approximately a one-to-one ratio between correct error detections and incorrect. However, four of the thirty-seven participants accounted for seventeen (nearly half) of the thirty-five false alarms. In the Form-Oriented task, there were a total of 111 false alarms (roughly a three-fold increase overall; maintaining a one-to-one ratio of error detections to false alarms). The mean false alarms per participant was 3.00 in the Form-Oriented task.

Over eighty percent of the participants increased their false alarms by at least one in going from the Meaning-Oriented to the Form-Oriented listening task. This fact is important in that it lends some support to the designations "meaning-oriented" and "form-oriented". It appears that most participants, with a few exceptions, were trying harder to consciously detect errors in the second task than in the first, and hence guessed more often in the second task than in the first. In other words, it appears reasonable to believe that participants tended to respond to errors that truly stood out to them in the first task, i.e., mainly concentrated on meaning, as intended. This is consistent with frequent comments of the sort, "I couldn't even think about whether there were any mistakes when I had to find the picture." To the extent that incorrect guesses indicate a conscious effort to find errors, it would appear that a small number of participants adopted a more strongly form-oriented strategy than most other participants during the Meaning-Oriented task. Overall, however, the first task appears to have been less form-oriented than the second, as intended. This only means that there was more effort to focus on form during the Form-Oriented task than during the Meaning-Oriented task. It does not prove that the improved performance in the Form-Oriented task was due to the greater attention to form, since other factors (more time, no secondary task, repetition of text) probably made

overall processing easier, which in turn could have increased the likelihood of more cues being processed and thus of incompatible cues clashing.

Given the number of incorrect guesses (participants were roughly half as likely to incorrectly mark a non-error-containing item as they were to correctly mark an error-containing item), it is highly likely that some of the tabulated "correct" responses were in fact not genuine error detections. This would decrease the systematicity of the distribution of reactions to the various. errors, making it harder to detect true differences between hypothesized error types.[17]

3.1.2.4. Printed Sentences

In the Printed Sentences task, participants reacted to an average of 5.89 (65.4%) errors per subject (range 2 to 8, $SD = 2.196$). The mean number of error detections in the Intermediate group was 4.88 (61%) out of the eight possible error detections per participant (ranging from 2 to 8; $SD = 1.96$). For the Advanced participants the mean was 6.75 (84.4%) of the errors (ranging from 2 to 8 also; $SD = 2.048$). The difference between the two groups in the Printed Sentences task was significant ($t(35) = 2.816$; $p < .01$, 2-tail). Sixteen of the twenty Advanced participants discovered seven or eight of the eight errors, as did four of the seventeen Intermediate participants. Only three of the twenty Advanced participants discovered two to three errors (and none less than two), while five of the seventeen Intermediates discovered two or three (and none less than two).

3.1.2.5. Printed Sentences Task and Listening Tasks

The Printed Sentences task in part provides a measure of off-line, metalinguistic knowledge. That is not to say that fluent readers might not react to the printed errors as cue clashes in an on-line, processing-related, automatic manner. However, since all of the participants had formal Russian training, and the errors involved basic uses of inflections, it is to be expected that participants would in general be able to find and correct errors on the basis of metalinguistic knowledge, even if they did not react to errors on an automatic, language-processing basis. Thus if scores in either of the two listening tasks were to be unrelated to scores in the Printed Sentences task, this would suggest an inability to apply metalinguistic knowledge in that other task. In the case of the Intermediate group, there was no significant relationship between the Printed Sentences task and the other tasks (for the Printed Sentences and the Meaning-Oriented tasks, $r = .23$, $df = 16$, $p > .05$; for the Printed Sentences and the Form-Oriented task $r = .18$, $df = 16$, $p > .05$). In the case of the Advanced group there was a relationship between the Printed Sentences task and both of the listening tasks (for the Printed Sentences task and Meaning-Oriented task, $r = .4432$, $df = 19$, $p < .05$; for the Printed Sentences and Form-Oriented Listening task, $r = .6712$, $df = 19$, $p < .01$).

[17] That is, highly detectable errors would yield scores closer to their true score (providing less opportunity for guessing) than would errors with low detectability, and therefore, the difference between items with high detectability and those with low detectability would be decreased as a result, making such differences more difficult to detect.

3.1.2.6. Error types and error detection

The instances of error detection were not randomly distributed among the errors, nor among the hypothesized error types. Tables 2 to 6 summarize these results for the three experimental tasks.

	Meaning-oriented Task		Form-oriented Task		Printed Sentences Task	
Error Type:	Inter-mediate group	Advanced group	Inter-mediate group	Advanced group	Inter-mediate group	Advanced group
Nominative (chelovek) replaced by instrumental (chelovekom)	1 (5.9%)	5 (25%)	8 (47%)	13 (65%)	15 (88.2%)	20 (100%)
Nominative (br'uki) replaced by locative (br'ukax)	0 (0%)	4 (20%)	7 (41.2%)	9 (45%)	10 (58.8%)	15 (75%)
Locative (krovati) replaced by instrumental (krovatju)	2 (11.8%)	6 (20%)	5 (29.4%)	10 (50%)	12 (70.6%)	14 (70%)
Instrumental (mylom) replaced by locative (myle)	0 (0%)	1 (5%)	3 (17.6%)	7 (35%)	11 (64.7%)	17 (85%)
Locative (kresle) replaced by nominative (kreslo)	1 (5.9%)	2 (10%)	3 (17.6%)	5 (25%)	10 (58.8%)	13 (65%)
Instrumental (polotentsem) replaced by nominative (polotentse)	1 (5.9%)	1 (5%)	2 (11.8%)	4 (20%)	12 (29.4%)	16 (80%)
Perfective (zakryl) replaced by imperfective (zakryval)	1 (5.9%)	8 (40%)	5 (29.4%)	13 (65%)	10 (58.8%)	14 (70%)
Imperfective (stojali) replaced by perfective (postojali)	0 (0%)	2 (10%)	3 (17.6%)	9 (45%)	9 (52.9%)	13 (65%)

Table 2. Errors detected in Experiment 1. In each Intermediate group cell the maximum number of possible detections was always seventeen and in each Advanced group cell the maximum possible was twenty.

Comparing the performance of the two groups on the various error types, one finds a rank order correlation between the two groups (Spearman ρ, corrected for ties) of .488 for the Meaning-Oriented listening task, which is not significant ($p = .196$) and significant rank order correlation of .826 for the Form-Oriented listening task ($p = .029$). However, the floor effect with the Intermediate group in the Meaning-Oriented task makes it impossible to establish a meaningful rank order for that task. Nevertheless, when we combine the results of the two groups for the Meaning-Oriented task, provided we

consider verb inflection and noun inflection separately, the rank order remains identical to that of the Advanced group alone. Likewise, if we combine the groups in the Form-Oriented task, the rank order of the Advanced group by itself is maintained except for the reversal in the order of the locative-for-nominative substitution (*br'ukax* for *br'uki*) and the instrumental-for-locative substitution (*krovatju* for *krovati*), which in either case are separated by a single error detection. In other words, it seems reasonable and perhaps desirable to combine the results of the groups in considering the distribution of errors and we shall proceed to consider the combined groups in what follows.

Type of Substitution	Number of participants detecting error	Phonetic salience of difference between error form and base-form (ignoring prefix)	Rank order predicted based on case function
Imperfective (zakryval) replacing perfective (zakryl)	9	additional syllable	n/a
Instrumental (krovatju) replacing locative (krovati)	8	additional syllable	mid
Instrumental (chelovekom) replacing nominative (chelovek)	6	additional syllable	high
Locative (br'ukax) replacing nominative (br'uki)	4	additional segment, vowel change	high
Nominative (kreslo) replacing locative (kresle)	3	no difference	low
Nominative (polotentse) replacing instrumental (polotentsem)	2	no difference	low
Perfective (postojali) replacing imperfective (stojali)	2	no difference	n/a
Locative (myle) replacing instrumental (mylom)	1	same length, vowel change	mid

Table 3. Experiment 1: Rank ordering of inflectional error detections in the Meaning-Oriented Listening task, combining the Intermediate and Advanced groups.

In considering the phonetic salience of the deviation of the erroneous form from the base-form, the perfective prefix is ignored, on the assumption that it is processed differently from suffixes. The difference being compared is that involved in word endings. The phonetic salience taken to be of interest is the amount of deviation of the erroneously inflected form from the base-form (i.e., the corresponding nominative form for nouns; the perfective form in the case of the imperfectivized verb and the imperfective form in the case of the perfectivized verb). Numerical values were assigned as follows: additional syllable—4; additional segment plus vowel change—3; vowel change only—2; no difference—1. Phonetic salience so characterized goes some way toward predicting the rank ordering of error detections in the Meaning-Oriented task ($\rho = .806$; $p = .033$). In order to compare the predictions based on case function (i.e., that oblique-for-

nominative substitutions would be noticed most, followed by oblique-for-oblique substitutions followed by nominative-for-oblique substitutions), with these predictions based on phonetic salience, we must limit ourselves to noun inflection. In doing that, the rank order correlation based on phonetic salience ceases to be significant ($\rho = .795$; $p = .076$). However, the correlation between the order predicted by case function and the observed order in the Meaning-Oriented task fares worse ($\rho = .478$; $p = .285$). (Note that it is in principle somewhat easier to achieve a significant rank order correlation with phonetic salience than with case-function since the former involves a four point scale and the latter a three point scale, while the error detections span a possible eight point scale.)

The rank ordering of error detections in the Form-Oriented task is given in Table 4.

Type of Substitution	Number of participants detecting error	Phonetic salience of difference between error form and base-form	Rank order predicted based on case function
Instrumental (chelovekom) replacing nominative (chelovek)	21	additional syllable	high
Imperfective (zakryval) replacing perfective (zakryl)	18	additional syllable	n/a
Locative (br'ukax) replacing nominative (br'uki)	16	additional segment, vowel change	high
Instrumental (krovatju) replacing locative (krovati)	15	additional syllable	mid
Perfective (postojali) replacing imperfective (stojali)	12	no difference (ignoring prefix)	n/a
Locative (myle) replacing instrumental (mylom)	10	same length, vowel change	mid
Nominative (kreslo) replacing locative (kresle)	8	no difference	low
Nominative replacing instrumental	5	no difference	low

Table 4. Experiment 1: rank order of error detections in the Form-Oriented listening task, combining the Intermediate and Advanced groups.

The rank order correlation for phonetic salience and error detections in the Form-Oriented task, following the same procedure outlined above, is somewhat stronger than in the Meaning-Oriented task and significant at the .05 level ($\rho = .883$; $p = .0487$). However, for the case-based prediction the rank order correlation is now even stronger ($\rho = .956$; $p = .033$) than for the predictions based on phonetic salience.

Since it is expected that the results from the Meaning-Oriented task would tend to be carried into the Form-Oriented task, it may also be worthwhile to consider separately just those error detections that occur for the first time in the Form-Oriented task. In comparing the number of new error detections with those of the Meaning-Oriented task, the correlation is relatively small ($r = .327$; $p < .05$). As with the total results from the Form-Oriented task, there is a strong rank order correlation between new error detections and case-based prediction ($\rho = .95$; $p = .033$) and a weaker, nonsignificant correlation with the phonetic salience based prediction ($\rho = .736$; $p = .1$).

Table 5 collapses the case errors into the three broader theoretical categories hypothesized, for both the Meaning-Oriented and Form-Oriented tasks.

Type of Substitution:	Error detections in Meaning-Oriented Task	Error detections in Form-Oriented Task	New error detections in Form-Oriented task	Order predicted based on case function
Oblique replacing nominative	10 (13.5%)	37 (50%)	27 (36.5%)	high
Oblique replacing oblique	9 (12.2%)	25 (33.8%)	16 (21.6%)	mid
Nominative replacing oblique	5 (6.8%)	14 (18.9%)	9 (12.2%)	low
χ^2 value	χ^2 (2) = 1.75, p = .417	χ^2 (2) = 11.89, p < .01	χ^2(2) = 9.519, p < .01	

Table 5. Experiment 1: Detections of case errors by general type of substitution, combining the Intermediate and Advanced groups.

Given the null hypothesis of equally distributed error detections among the types, we find χ^2 values for the observed distributions to be 1.75 ($p = .417$) for the Meaning-Oriented task and 11.891 ($p < .01$) for the Form-Oriented task as a whole and 9.519 ($p < .01$) for the new items detected in the Form-Oriented task. That the Form-Oriented task appears to reveal a systematic pattern that is not detected in the Meaning-Oriented task raises the suspicion that the so-called Form-Oriented task is revealing the effects of genuine language processes which were too slow or inefficient to function well in the Meaning-Oriented, task and which were statistically overwhelmed by the effect of metalinguistic knowledge in the Printed Sentences task (see below). When we collapse the errors according to types as we have done here, the differences in phonetic salience between the oblique-for-nominative and oblique-for-oblique substitutions largely cancel out, while the observed order of frequency of detections is the order predicted by the case-based considerations.[18]

Another important point in relation to the relative ability of phonetic salience and case function to predict the distribution of error detections is the impact of the instrumental-for-locative substitution in these data. If it is removed, the picture changes. Just considering the Form-Oriented task, without the instrumental-for-locative substitution, the rank order correlation between phonetic salience and observed error detections is $\rho = .718$, $p = .151$, while for the case-based predictions the correlation is $\rho = .949$, $p = .100$. Thus we can largely suspect this one error as accounting for the apparent advantage of phonetic salience over case-function as a predictor of sensitivity to errors, while its high detectability may in fact be due to factors other than phonetic salience (see below). And as we have seen, when the error types are collapsed, the distribution of errors relates well to case function even when phonetic salience differences are largely equalized.

Having separated case-errors and aspect-errors for rank ordering, we are left with only two aspect errors. The distribution of errors appears to be non-random ($\chi^2[1] = 5.143$, $p < .05$, for the Meaning-Oriented task and $\chi^2[1] = 39.364$, $p < .001$ for the Form-Oriented task). Recall that there was a question over the independence of as many as six

[18] It is perhaps worth noting that an analysis of variance with the collapsed error categories as the independent variable shows an effect of the case-based variable on error detections in the Form-Focused task ($F[2,3] = 7.359$) which approaches significance ($p = .070$). If we limit ourselves to only the items noticed first in the Form-Focused task, the effect ($F[2,3] = 17.643$) is significant ($p = .22$). When phonetic salience was treated as the independent variable, no such effects were found.

of the Advanced group responses to the imperfective-for-perfective substitution in the Meaning-Oriented task. This somewhat undermines confidence regarding the difference in detectability between the two aspect errors. However, the increase from 1 to 5 detections by the Intermediate group in going from the Meaning-Oriented task to the Form-Oriented task could conceivably restore a bit of confidence, since there was no problem with the independence of responses for that group.

Turning to the rank ordering of error detections in the Printed Sentences task (combined groups), we find that any statistically significant systematicity vanishes. The rank ordering of errors detected in this task does not correlate with either the Meaning-Oriented listening task ($\rho = .0422$, $p = .911$) nor the Form-Oriented listening task ($\rho = .144$, $p = .738$). Likewise the χ^2 test does not allow us to reject the null hypothesis of an equal distribution of error detections across errors ($\chi^2[7] = 7$, $p = .429$). Collapsing errors across types rather than looking at individual errors, which improved the appearance of systematicity in the Form-Oriented task, does not help in this case ($\chi^2[7] = .6296$, $p = .730$). In other words, the effect of the participants' possession of the relevant metalinguistic knowledge appears to be distributed randomly over error types, while the factors that affect error detection during on-line listening comprehension appear to be systematic, especially in the so-called Form-Oriented task. The rank ordering of error detections for the Printed Sentences task is given in Table 6:

Instrumental replacing nominative	36 detections
Locative replacing instrumental	28 detection
Nominative replacing instrumental	28 detections
Instrumental replacing locative	26 detections
Locative replacing nominative	25 detections
Perfective replacing imperfective	24 detections
Nominative replacing locative	23 detections
Imperfective replacing perfective	22 detections

Table 6. Experiment 1: rank ordering of error detections in the Printed Sentences task, combing the Intermediate and Advanced groups.

3.1.3. Discussion

The first prediction we set out to test was that **L2 Russian users with greater experience in the language will show more sensitivity to inflectional cue-clashes than those with less experience**. The difference between the two groups is clear. Some serious learning related to inflectional morphology (along with a general increase in ease of processing) is presumably behind the advantage of the Advanced group over the Intermediate groups in the listening tasks. The pace of development appears to be slow. My consultant, Tamara Alexandrovna Ivanova, expressed strong surprise that participants could miss so many basic, striking errors. This finding clearly contradicted what she had expected based on participants' typical performance in spoken and written production in her classroom. Nevertheless, our findings suggest that learning is taking place.

The second prediction was that there would be **systematic patterning in the tendency of participants to notice some error types earlier or more often than others**. The systematic patterning did not appear in the Meaning-Oriented task, arguably because, overall, not enough errors were detected in that task for a pattern to emerge. In

the Printed sentences task there was also no significant pattern, suggesting that the systematic tendencies that might have resulted from on-line processing were statistically swamped by the effects of unbridled application of metalinguistic knowledge. Thus the systematic data appeared only in the Form-Oriented task, where, I would argue, there were enough error detections for a pattern to be detectable, but not enough effects of metalinguistic analysis to hide systematic effects of spontaneous language processing mechanisms. The nature of the pattern of sensitivity to errors will be discussed at length below.

The third prediction was that **error detections would increase with increasing opportunities to apply metalinguistic analytical strategies.** This prediction also appears to have been supported. The participants had two opportunities to detect errors while listening. On the second listening, they were listening to a story that they had originally become familiar with not only from listening, but from listening with considerable pictorial support. The sentences were all identical in form to those heard before, as well. And yet, when in the third task participants were allowed to freely examine the sentences and consciously reflect on their grammatical, they were able to nearly double the number of errors they discovered. (Mean errors detected went from 3.027 in the second listening task to 5.892 in the Printed Sentences task.)

We next turn to various issues related to the three tasks and their interconnections, and then go on to consider a variety of other implications of the findings. We also more carefully examine the individual Russian errors that were employed in the experiment.

3.1.3.1. Meaning-Oriented Listening

The Meaning-Oriented task allowed little scope for metalinguistic reflection, but ample scope for reaction to errors if they were to raise a problem in on-line processing, that is, if participants were to experience cue clashes. In this task, Intermediate participants showed almost no sensitivity to errors. Advanced participants showed sensitivity to nearly a quarter of the errors on average, with a few reacting to over fifty percent of the errors. This task was intended to provide a window into the extent to which inflection plays a role in normal communication, when there is no special attention to form. Keep in mind that native control listeners were strongly affected by the errors, in spite of the fact that, like the experimental participants, they too were required to concentrate on the meaning of the sentences. The experimental findings suggest that although inflection is not processed very much at all during the stage of learning represented by the Intermediate group, enough processing of inflection must go on in their lives to account for learning reflected in the increased sensitivity seen in the Advanced group, given the assumption that a pseudolongitudinal experimental design can reveal developmental differences (Gass & Selinker, 1994). As a matter of fact, it is difficult to imagine any explanation for the difference between the two groups other than developmental advances. When opportunity for metalinguistic reflection is virtually eliminated (the text segments following rapidly one after another, and a secondary task keeping attention strongly focussed on meaning), learners differ, based on how long they have been learning/using the language, in terms of their tendency to react in a nativelike way to inflectional errors. This is a necessary (though perhaps not sufficient) finding if we want to claim that from the standpoint of listening comprehension, the acquisition of inflectional processing is under way in such learners.

3.1.3.2. Form-Oriented Listening

The Form-Oriented task was different from the Meaning-Oriented task in two major ways. First, it provided less demanding conditions for processing (the texts are already familiar, there is no secondary picture selection task, and the time for processing is increased). Second, attention to form was explicitly encouraged. It was no longer a matter of students noting errors that jumped out at them. Rather, they were instructed to deliberately search for errors.

Both groups show a substantial increase in error detections on the Form-Oriented task as compared to the Meaning-Oriented task, suggesting that factors such as the ones just noted can probably affect the likelihood of inflection being processed. To a reasonable extent, if a participant reacted to errors in the Meaning-Oriented task, an increased score in the Form-Oriented task was likely, since the participant would be adding to the errors already noticed during the first task. However, matters turned out not to be so simple. The Intermediate participants as a group displayed extremely low sensitivity to errors in the Meaning-Oriented task, while as noted, in the Form-Oriented task, their number of reactions, although still relatively low, showed a five-fold increase. For the combined groups there was a reasonable correlation between scores in the two listening tasks ($r = .589$, $r^2 = .347$, $p < .05$), probably reflecting the expected effect of error detections in the first task being carried over into the second. However, there was far from a simple additive relationship. There appears to be no relationship between the tendency to react to errors in the first task in relationship to the tendency to detect *additional* errors in the second task ($r = .055$). That is, some participants had already reacted to most of the errors they would detect in the Form-Oriented task during the Meaning-oriented task, while for others, most of the errors noticed in the Form-Oriented task were noticed there for the first time. The difference in performance on the two tasks was significant overall ($t[36] = 7.1$, $p < .001$). However, it appears that in a sense, the Form-Oriented task may have uncovered an approximate maximum ability of an individual to detect the particular errors in spoken texts. Some participants approached this maximum in the first task, in spite of the heavy processing demands. Others could only approach it under the decreased processing demands of the second task. The Intermediate group fell entirely within the latter category, as did some members of the Advanced group. That is, the groups differed overall in terms of their error detections in the combined listening tasks, but within the advanced group there appears to have been a further difference between subjects based on susceptibility to effects of increased/decreased processing demands.

Although the evidence from the increased incidence of false alarms in going from the Form-Oriented to the Meaning-Oriented tasks suggests that participants attempted to be more form-focussed during the former than the latter, in fact, it remains possible that error detections in the former may have been more due to on-line cue clashes than to metalinguistic reflection. For one thing, it would be virtually impossible to detect an error without also understanding the sentence. The question would then be whether, after understanding the sentence, the listener would be able to hold it in short term memory for purposes of deliberate analysis. I should point out that the second listening task replaced an attempted task in an earlier experiment which was aimed at detecting effects of deep versus shallow processing on error detection. It seems possible that in the present experiment what the key variable distinguishing the two listening tasks was processing demands rather than opportunity to apply analytical knowledge. If so, rather than a meaning-oriented listening task and a form-oriented listening task, we may have ended up with two meaning-oriented tasks (i.e., in which the mental resources were depleted by comprehension processes, with little further possibility of applying metalinguistic

knowledge in an analytical, problem-solving mode). This possibility will be discussed further below. If it is correct, then for many participants, on-line processing was strongly sensitive to overall processing demands. That would leave the Printed Sentences task as possibly the only task with a high potential for eliciting error detections that were based on conscious application of metalinguistic knowledge about grammatical form.[19]

3.1.3.3. Printed Sentences Task

The Printed Sentences task was intended to allow maximum exploitation of metalinguistic analysis in detecting errors. It also provided some control over the genuineness of the error detections in the other two tasks, since in this task, the participants were asked to indicate the location of the error and to correct it. Again, it was expected that there would likely be a natural increase in scores based on the fact that errors detected in the first two tasks would be carried into this task. However, it was felt that this would primarily affect the speed with which learners performed this task. Since there was no time restriction in this task, participants were free to go over the sentences as many times as they wished. The aim was to determine the level of error detection when maximum scope was allowed for applying metalinguistic knowledge to the analysis of the sentences containing the errors. To the extent that inflectional processing mechanisms in reading comprehension have developed for a participant, reactions here could again be due to cue-clashes. However, reactions were clearly not limited to those arising from L2 processing mechanisms. If need be, participants could apply their metalinguistic analytical knowledge in a nonlinguistic, problem-solving mode. Thus the task allowed complete scope for the employment of so-called explicit grammatical knowledge.

With the new possibility of heavy off-line exploitation of metalinguistic analysis, almost all participants were able to identify errors that they did not detect, even with the benefit of two opportunities, when listening to the same stimuli. Nonetheless, it is quite interesting that the Intermediate participants found less than half of the errors even under these conditions, while the Advanced participants detected over eighty-four percent of them. In fact, most, if not all, of the Intermediate learners had already received the relevant formal grammatical teaching. (Recall that from the standpoint of Russian grammar classes, the errors in the stimuli involved inflectional forms and uses that are covered even in very basic grammatical training.) Typically, grammar exercises in formal language classes in Russia focus on series of very narrow grammar topics. I have observed that students at the level of our Intermediate participants are often able to display their metalinguistic understanding related to the point of the day. It may be that the mastery of this knowledge is sometimes less secure than intended by teachers, and becomes more secure only as overall proficiency increases. That makes sense in that increased proficiency might increase the likelihood of there being some new linguistic categories inside the learner that a given bit of metalinguistic terminology might refer to, making it more possible to grasp the meaning of the terminology and retain that bit of metalinguistic knowledge.

The widely debated issue of the relationship of formal instruction to L2 development is typically explored by examining the effects of variables such as "explicit

---

[19] In the broader sense, all three tasks are metalinguistic in that learners had to make error detection decisions. However, the first task (if not the first two tasks) depends on on-line, automatic reactions to errors. Listeners could react to an error if it caused a processing glitch, but could not reflect on or analyze the sentences. The third task allowed free reflection and analysis. The difference between the results in first task and the third task are enormous, consistent with the idea that they tapped different cognitive processes.

grammar teaching" or "explicit correction" on changes in the learners linguistic behaviour (e.g., Ellis, 1990; Doughty, 1991, Lightbown & Spada, 1990; Spada & Lightbown, 1993; White, Spada, Ligtbown, & Ranta, Leila, 1991; Lyster & Ranta, 1997; Spada, 1997). The role of instruction in relation to specific features of target languages is sometimes considered in relation to the developmental readiness of learners for those features (Meisel, Clahsen & Pienemann, 1981; Pienemann, 1985; Lightbown, 1998). Lightbown (1998) argues that presenting grammatical forms that are developmentally advanced for particular learners may be of value. For example, it may lead to immediate increases in the number of nativelike utterances, even if these are based on memorized chunks or "monitoring". Furthermore, the improvements seen in Spada & Lightbown (1993) continued to increase for some period after training (unlike the improvements seen in White, 1991). Lightbown (1998) suggests that explicit instruction may have planted the seeds for later development. Such studies do not typically attempt to examine changes in comprehension or production mechanisms that may or may not result from the particular treatments. For example, it would be interesting to know whether after learners receive formal instruction related to L2 English question formation, sentence initial auxiliary position begins to function for those learners as an on-line cue to the interrogativity of English sentences, or whether that particular cue is as inert after training as it was before. In the case of (pedagogically elementary uses of) Russian inflectional morphology, does the typical training primarily affect processes such as those involved in the online cue-clash detection of the Meaning-Oriented task? Does it mainly affect the additional processes involved in the off-line Printed Sentences task? Unlike English question formation (Spada & Lightbown, 1993), but perhaps more like English adverb placement (White, 1991), it may be that for many learners neither the on-line nor off-line skills are immediately and strongly susceptible to pedagogical intervention, but rather require some period of continued learning. As was suggested earlier, there may also be a partial dependency of the development of either or both of these types of processes on the development of the other. In particular, the ability to retain and apply the relevant metalinguistic, analytical (explicit grammatical) knowledge may be affected by the development of genuine processing mechanisms. In short, it would be of value to ask what processes and abilities are behind the effects of a pedagogical treatments. Examination of production data, or data from off-line tasks, may not distinguish between the metalinguistic, linguistic, and other cognitive components of performance changes brought about by instruction or correction.

As far as the statistical relationship of the Printed Sentences task to the Form-Oriented task is concerned, in the case of the Advanced group the correlation ($r = .653$ $p < .01$) is not impressively large, since, by design, errors detected in the Form-Oriented task constituted a substantial subset of those detected in the Printed sentences task. In other words, given enough error detections in the former, there was bound to be a reasonable correlation with the latter. The absence of a significant correlation ($r = .18$, $r^2 = .036$, $df = 15$, $p > .05$) between these two tasks in the case of the Intermediate participants is more interesting. It suggests that the metalinguistic knowledge (tapped by the Printed Sentences task, in which Intermediate scores were fairly evenly distributed over the whole range of possible scores) played little or no role in the five-fold improvement exhibited by the Intermediate participants in going from the Meaning-Oriented to the Form-Oriented task (where scores ranged from 0% to 50%). This supports the suggestion that normal language processing mechanisms account for much of the improvement seen in going from the Meaning-Oriented task to the Form-Oriented task. For this reason, the Form-Oriented task has occasionally been referred to as the "so-called Form-Oriented task". Yes, participants made a greater effort to focus on form in the second task (as evidenced by the incidence of incorrect guesses); however, their

performance still relied heavily on their normal (meaning-oriented) language processing mechanisms. In fact, before a listener could successfully apply metalinguistic knowledge to the analysis of an utterance, s/he would have to understand the utterance. To know, for example, that a grammatical subject should not be marked with instrumental case, it is necessary to understand that the noun is the agent (etc.) and not the instrument (etc.). By contrast, the Printed Sentences task allows more opportunity for applying metalinguistic problem-solving strategies, so that in that task, someone with no comprehension or production proficiency at all, but adequate metalinguistic knowledge, might find the errors. It is not reasonable to imagine that such an individual could hold a sentence in verbatim auditory form in working memory while carrying out such metalinguistic problem-solving. For that, it would seem that the stability of the printed form is essential.

3.1.3.4. On the significance of reactions to inflectional errors

The presence of nativelike reactions to cue clashes (in this case, inflectional errors) in L2 learners would be a necessary, though not sufficient, condition for attributing nativelike comprehension processes to them (suggesting, for example, syntactic as opposed to purely semantic processing). It might be helpful to keep in mind the possible mechanisms behind the reactions of the native control participants. The native mental language processor is receiving information from multiple sources (lexical cues, grammatical cues, the existing discourse model, knowledge of the world) and expects this information to converge. A failure of convergence may indicate a processing error. In Chapter 1, we noted that there is considerable evidence that meanings are constructed on-line during sentence comprehension, although there may be controversy over the points at which various information sources become integrated (Rayner, Carlson & Frazier, 1983; Trueswell, Tanenhaus, & Kello, 1993; Altmann, 1999). Earlier features of a sentence create expectations regarding the later portions of the sentence. For example, when a Russian native speaker hears a sentence that begins with a noun phrase marked with instrumental case, the possible continuations of the sentence are restricted. The possibilities will include a passive sentence continuation (in which case the instrumental noun phrase will turn out to be an overt agent of the passive verb) and a predicate noun construction (in which case the instrumental noun phrase will turn out to be a preposed predicate noun phrase). In one of our experimental sentences, a subject noun (requiring nominative case) was incorrectly inflected with the instrumental case. This sentence is given below as (1). In processing (1), at some point, perhaps when the active verb *myt'*, 'wash' is encountered, the native mental language processor would presumably start to encounter difficulty, since there is no available role in the clause for an instrumental case-marked human noun. By the time the end of the sentence is reached, the initial noun phrase will not have found any role in the sentence compatible with its case marking, and the listener will be have experienced the error as an error (a failure of cues to converge).

(1)	Potom	*chelovek-om	stal	myt'	ruki	pod	kranom.
	then	person-instr	started	to.wash	hands	under	tap.

Then the man started to wash his hands under the tap.

In contrast with this situation, where the case marking would set up expectations that subsequently go unfulfilled, there are cases such as (2), also one of the experimental sentences, in which the native processor might be expecting a case marking that does not materialize.

(2)	Nakonets	chelovek	vyter	ruki	*polotentse.
	Finally	person:nom	wipe	hands	towel:nom/acc

Last of all, the man wiped his hands with the towel.

Before the word *polotentse* 'towel:nominative/accusative' is completely uttered, the native listener may have identified the target word as *polotentsem*, 'towel:instrumental' (Marslen-Wilson & Tyler, 1980). That is, we saw evidence in Chapter 1, in the discussion of the Cohort Model of lexical access, that a word would typically be identified at the point at which it is distinguishable from all other words. The remainder of the word-form will then need to be compatible with not only lexical, but also morphosyntactic and semantic expectations. Rather than setting up subsequently unfulfilled expectations as with example (1), in example (2) the case-marking itself might violate a strong expectation that was already set up by the earlier part of the sentence.

Either of these types of violations of expectations appears to trigger a strong reaction in native speakers. This was qualitatively evident from observation of the native controls as they participated in the experiment. In many cases participants were observed immediately marking the "X" (indicating an error) on their answer sheet, and then subsequently searching for the correct picture. Errors of the first sort (assigning an early interpretation that proves impossible later in the sentence) might be expected to create a more consistent reaction than errors of the second type (strongly constrained forms not occurring where required). In example (1), the phonetic form of the incorrectly inflected noun will not yet have other information sources competing with it and can only be heard as a noun in the instrumental case. With sentence (2), it is possible that the expectation of the instrumental ending would be strong enough to override the actual form occurring, with the native listener mentally "restoring" (Samuel, 1981; Marslen-Wilson, 1973) the ending, especially when the mistaken and correct endings are both short and unstressed. Such a fluent restoration was observed during preparations for the abandoned experiment mentioned earlier, such that a stimulus sentence had to be radically altered because of the fact that an error that one of my native Russian consultants quite deliberately pronounced appeared to be hardly detectable to my second native Russian consultant, although the substitutions involved a difference in vowel quality ([o] versus [u]).

As noted, if the nonnative experimental participants were to react similarly to native controls when processing sentences containing errors, that would be prima facie evidence that similar processing mechanisms have developed (assuming we can prevent the free application of metalinguistic knowledge thus forcing listeners to detect errors as a result of on-line comprehension processes). This is not what we found. Consider sentence (1), for example. During the Meaning-Oriented task, whereas 100% of the native control participants reacted to the error, only sixteen percent of the experimental participants reacted to it (twenty-five percent of the Advanced group and 5.9% of the Intermediate group). There were many factors potentially favouring the detection of the wrong ending by experimental participants. The word *chelovek*, "person" is a high frequency word (ranked number 63 in Brown's [1996] list of 10,000 high frequency Russian words). It had recently occurred in the same text in the nominative form. Thus we might consider the base-form to be primed, increasing the likelihood that the affixed form could stand in clear contrast with the base-form, allowing the instrumental case form to trigger the relevant components of the comprehension process, leading to a cue clash. The remaining words in the sentence are also common words, and the entire event described is highly predictable (the man had just been said to have gone into the bathroom because his hands were dirty). Thus, in various respects this should qualify as a relatively easy sentence to process and to integrate into the discourse model. Furthermore,

the incorrect affix adds a fourth syllable to the high frequency three syllable base-form. Yet the case form had no detected effect on the way in which the sentence was processed by seventy-five percent of the Advanced participants, many of whom have lived in Russian and functioned as university students for some time.

Overall, 11.5% of the errors were noticed by the experimental participants (18.1% by the Advanced participants) in the Meaning-Oriented task. This contrasts with ninety-eight percent of the errors being noticed by native controls performing the same task.

3.1.3.5. The pace of development

Certain cautions are important in interpreting these results. The presence of inflectional errors may have made nonnative performance appear less nativelike than it would be without the errors. That is, in the face of complications, the nonnative inflectional processing system might tend to shut down, while working much better during smooth sailing. Kilborn (1991) reported that German speaking participants, who normally relied substantially on case in assigning grammatical relations, switched to a greater reliance on word order under noisy conditions. There is no reason to think that L2 Russian speakers in whom a case processing system had partly developed would not at times revert to, say, a stronger reliance on word order (especially where word order is a powerful cue in the L1 of a participant). As a matter of fact, given the extended time period over which the nonnative system appears to be developing (if in fact it is developing), nonnative speakers are presumably well-practiced at disregarding the details of inflectional form that they cannot yet make use of. Meaningful processing requires minimally that lexemes be identified and used for mental model construction. For example, a nonnative speaker confronted with *chelovekom*, 'person:instrumental' in the sentence *Moja sobaka xotit stat' chelovekom*, meaning 'my dog wants to become a person', might determine the function of the word *chelovekom* as a predicate noun by relying on a word order strategy, while ignoring case form. This might be necessary even once instrumental case has become active for the learner in its more basic function as a marker of instrument nominals. That is, even when more basic functions are acquired, case cues presumably remain unusable for less basic functions. If the nonnative listener is to understand the sentence however, s/he cannot get around the need to identify the word itself as a form of the lexeme to which it belongs. This necessary minimum might be accomplished through a prototype organization of the lexicon with the nominative form, *chelovek* as the prototype and other inflected forms treated as exemplars of it (see Chapter 4). As time goes on, for some functions, the instrumental-marked word-form might come to trigger additional processing beyond lexeme identification, but in other contexts it will simply get in the way if not suppressed. Thus participants' failure to react to a wrong case-marking in a given context under our experimental conditions could coexist with a tendency to process the *correct* case forms more normally in similar contexts. Analogous comments would apply to aspect inflection.

Another caution goes in the other direction. There is no guarantee that increased ability to detect errors is entirely due to increased use of inflectional forms for language processing. Many aspects of the overall comprehension system will presumably become more efficient and less resource-consuming over time. This could make it increasingly possible over time for L2 listeners to detect case errors metalinguistically. Getting around this problem entirely would require a more implicit approach to detecting sensitivity to errors than the approach taken here (e.g., a word-monitoring task with targets located after errors vs. after non-errors). In the present experiment it is safest to assume that the number of detected errors is probably somewhat greater than what we might find in an implicit task.

Even with such cautions in mind, the contrast in the Meaning-Oriented task between the floor effect with the Intermediate group (an average of .35 error detections per participant, out of eight possible detections) and the ceiling effect with the native controls (average of 7.8 errors detected per participant, out of the same possible eight), provides a perspective in which to consider the possible progression toward nativelike performance on the part of the Advanced participants. The progression, though statistically significant, is surprisingly small (an increase of 1.5 in the average number of error detections per participant). That is in fact a four-fold increase and represents an additional 18.8% of the errors. Yet the difference between experimental participants and native controls remains the more dramatic fact. Recall that the Intermediate participants had been learning Russian on average for over two years, while the Advanced participants had been learning Russian on average for over five years. If what we were expecting to see was steady, rapid progression toward a fully robust, nativelike inflectional processing system, these results clearly contradict our expectations.

In the two texts processed by the participants in this experiment, there were 296 opportunities for the thirty-seven individual participants to react to the eight tabulated inflectional errors. There were in fact only 211 error detections in the Printed Sentences task, which we might take to mean that eighty-five possibilities were thus out of the running in the listening tasks. Of these 211 remaining possible reactions to contextually impossible inflections during the Meaning-Oriented listening task, there were only thirty-five instances of participants reacting. Only six of these reactions were by members of the Intermediate group (who have been learning Russian for an average of 2.15 years, living in Russia an average of 1.65 years, using Russian regularly for a variety of purposes). The other twenty-nine instances of participants reacting to errors in the Meaning-Oriented task were by members of the Advanced group (learning Russian an average of 5.75 years, living in Russia an average of 3.46 years). This unexpectedly low incidence of reactions to inflectional errors occurred in spite of the facts that (1) the texts were not demanding (involving common vocabulary, relatively simple structures, concrete semantics and highly predictable content) and (2) the correct target inflections were limited to some of the most elementary functions of those inflections, functions which would have been covered early in formal training and reinforced for many participants repeatedly in exercises and through error-correction.

3.1.3.6. The Form-Oriented Task and inflectional processing

The Form-Oriented task must also have included a strong component of normal language processing, since the sentences would generally need to be understood in order to be judged erroneous. As noted earlier, in that task there were additional advantages that would presumably increase the overall ease of processing: (1) the identical sentences had been heard shortly before, with visual contextual support that would have perhaps made the narratives more memorable; (2) there were ten-second pauses between items, allowing additional time for end-of-sentence processes which might have been functioning too slowly to keep up with the pace of the first task. Presumably, the full range of comprehension processes, from speech perception to lexical access, to whatever syntactic and semantic processing is going on, including any implementation of inflectional morphology, are slower for nonnative speakers than for native speakers. The extra time available in the Form-Oriented task could allow increased utilization of processes that were in principle available to participants during the Meaning-Oriented task, but too slow for full functioning. Such a conclusion is supported by the pattern of error detection, which was generally similar in the Form-Oriented task and the Meaning-Oriented task (even when only new detections of the Form-Oriented task were taken into account),

while in the Printed Sentences task there was no such pattern, nor any significant systematicity whatsoever.

In addition to whatever effects there may have been due to genuine L2 comprehension mechanisms, we saw evidence from the number of false alarms that an effort was being made in the Form-Oriented task to exploit metalinguistic knowledge. Under these conditions, out of these 211 potential instances of individual participants reacting to errors (i.e., those we can consider in the running based on performance in the Printed Sentences task), there were 106 reactions (in the case of the Advanced group, out of 135 possible reactions there were seventy-eight). Metalinguistic knowledge, in view of the lack of systematicity in the responses in the Printed Sentences task, would most likely add "noise" to the data, rendering systematic effects of comprehension processes more difficult to detect. If so, then the apparent systematicity of the reactions in the Form-Oriented task may indicate that the results are due *mainly* to comprehension processes, with metalinguistic knowledge playing a relatively small role. That is, nonnative listening under decreased pressure may become more similar to native processing. Native comprehension processes are highly automatic and are thus not affected by the divided attention feature of the Meaning-Oriented task or by the steady pace of that task. Even if nonnative processing has taken on some of the features of native processing for many participants, it obviously lacks the speed and automaticity of native processing.

3.1.3.7. Problematic items

When we compare only the new error detections of the Form-Oriented task with the error detections in the Meaning-Oriented task, we find that the main difference in patterning involves the oblique-for-oblique substitutions. Reactions to the locative-for-instrumental substitution (*myle* for *mylom*) started out unexpectedly low in the Meaning-Oriented task (only one detection), while reactions to the instrumental-for-locative substitution (*krovatju* for *krovati*) started out surprisingly high in that task (eight detections). In the Form-Oriented task, there were nine new detections of the former and seven new detections of the latter (i.e., more of a balance). Neither of these increases was as large as the increases in the number of detections of oblique-for-nominative substitutions (increases of fifteen and twelve) nor as small as the increases in the detections of nominative-for-oblique substitutions (increases of five and four, respectively). It is possible that the behaviour of these two items is an artifact of the experiment. This possibility will be considered below under the discussion of individual items.

3.1.3.8. Metalinguistic knowledge and global proficiency

It was assumed that if participants had metalinguistic knowledge of the basic uses of the erroneous inflected forms and also of the correct target forms required by the contexts in which the errors occurred, this knowledge would be apparent in the results of the Printed Sentences task. The task was extremely straightforward in that only text segments with errors were included and participants were informed that there was one and only one error per text segment. Of course, a native Russian reader would presumably be able to perform well in this task by relying on normal language processing (reading) mechanisms without employing metalinguistic analytical knowledge. If a nonnative reader were responding on the basis of comprehension processes rather than metalinguistic knowledge, we might hope to see some systematicity in the responses, with certain errors more likely to be corrected than others. The difference in ability on this task between Intermediate and Advanced participants is interesting (sixty-one percent vs. eighty-four

percent of errors detected, respectively), in that the relevant metalinguistic knowledge is of the very basic sort that would have been covered relatively early in the formal courses in which the participants took part early in their Russian learning. The

Intermediate students' scores were distributed fairly evenly over the range of possible scores in the Printed Sentences task, each score from one to eight being represented by either two or three participants. The Advanced participants were skewed toward the high end, with eleven participants (fifty-five percent) scoring eight out of eight, and five more of them (twenty-five percent) scoring seven. Only three Advanced participants (fifteen percent) scored in the low range (one or two detections).

It is somewhat doubtful that many of the participants in either group were totally lacking in the relevant metalinguistic knowledge for any item. Given an even simpler task, such as one in which the erroneous words were typographically highlighted, we might find evidence of metalinguistic knowledge that went undetected in the present task. In any case, it remains clear that the ability of the Intermediate participants to apply such knowledge as an aid to language processing during listening comprehension is extremely limited, even for those with high scores in the Printed Sentences task. And to the extent that Intermediate participants were simply lacking in relevant metalinguistic knowledge, we would be led to conclude that despite the best intentions of language teachers, some learners do not master and retain in any usable form certain aspects of metalinguistic knowledge that are beyond their developmental level. Rather, for many participants, metalinguistic knowledge would appear to become more secure as overall proficiency becomes more advanced. And in any case, for intermediate participants with a greater tendency than average to acquire metalinguistic ("explicit grammatical") knowledge at an early stage, there is no commensurate tendency to process the particular features of linguistic form during listening comprehension.

In the Advanced group matters could be otherwise. Strength of performance in the Printed Sentences task significantly covaried with strength of performance in the listening tasks, with nearly twenty-percent shared variance in the case of the Meaning-Oriented task and forty-five percent shared variance in the case of the Form-Oriented task. However, given that metalinguistic knowledge appears not to have seriously increased the error variance in the Form-Oriented task (at least not enough to destroy the systematicity of the results in that task), we might still wonder to what extend the shared variance reflects an effect of metalinguistic knowledge on performance in the Form-Oriented task and to what extent the Printed Sentences task, like the other two tasks, benefited from improved L2 processing mechanisms. Such an improvement in processing mechanisms might also partly account for the difference in performance between the Intermediate and Advanced groups in the Printed Sentences task, for that matter.

3.1.3.9. Phonetic salience versus deeper processing explanations

I went to some lengths to point out the possible confounding of phonetic salience with deeper processing explanations, especially in connection with case-marking. However, the ability of phonetic salience to predict the distribution of error detections depended on certain theoretical commitments. In determining the degree of phonetic salience of a form, a researcher with a different theory might have compared oblique noun forms to unsuffixed stems, or contrasted erroneous forms with the correct target forms, or treated prefixes as contributing to salience as much as suffixes. Any of these alterations in the approach to salience could have yielded a different effect, or no effect, of phonetic salience.

In any case, it may well be the case that there is a sense in which being phonetically salient and being a trigger of particular comprehension process *should*

correlate. If an inflectional form differs from the base-form only by a difference of vowel quality in a final unstressed syllable, the lexeme itself may be identified quite readily. It would have been possible to use examples with more radical inflectional alternations than we used. For example, an inflectional alternation can involve the addition of two syllables, a loss of a final vowel with a corresponding reduction in the number of syllables, a change in the location of stress with concomitant alterations of vowel quality within the stem, the syncope or reduction of a stem-internal vowel, or the epenthetic insertion of a stem-internal vowel that might itself be stressed and qualitatively unlike any vowel in the base-form. To be concrete, consider the noun *m'ach*, 'ball' and imagine a child L1 learner or an adult L2 learner for whom this is a familiar lexical item in this base-form and thus accessed in this form from the mental lexicon rapidly and easily. Imagine on the other hand that the learner has never encountered the instrumental plural form *m'achami*. The phonetic form of the familiar base-form is ['mʲætʃ] while the instrumental plural phonetic form is [mʲɪ.tʃámi] (with the stem coda vowel reduced and stem final syllabic coda resyllabified, at least impressionistically, as the onset of the totally new stressed first syllable of the two syllable affix, the result being three syllables in contrast with the single syllable base-form). In such a case the listener's challenge in identifying the lexeme is substantial. That is, the mere process of finding the lexeme centered around the familiar base-form is going to draw heavy resources, in the process drawing extra attention (in some sense) to this word and to its specific inflectional form. In keeping with grammaticalization theory, in the course of diachronic language development, the more frequent, more predictable affixes will tend to become unstressed, to reduce and erode, while the less frequent and less predictable ones, i.e., the ones that *need* to attract more attentional resources, erode less (Bybee et al., 1994). Extra phonetic distortion would thus help to focus resource where needed for extra processing below the phonetic level. Thus the variables of phonetic deviation from base-form and the variable of deeper processing functions are likely to be closely related in general, causing us to question the logic of attempting to examine them separately. Rather than seeking to totally separate the variable of phonetic salience and deeper functional factors, we might have to be content with avoiding extreme differences in salience, as we have done for the most part in the current experiment, and provide evidence (see the qualitative discussion of individual items below) that all crucial forms are adequately perceptible. At this point, I take the position that the degrees of difference in phonetic salience involved in this experiment are too small within the full speech stream to provide a complete account of the systematicity in the distribution of reactions to errors.

3.1.3.10. Deeper processing considerations and case-errors

The low detectability of nominative-for-oblique substitutions would seem to require two explanatory factors: (1) the nominative case form is not triggering its own processing and (2) the processor is not looking for oblique marking in oblique contexts. It is important to understand these two variables, as they play a major role in the set of hypotheses that will be drawn from the results of Experiments 1 and 2. From the deeper processing perspective, the highest possible detectability ought to result from a combination of (1) the encountered case category triggering its own processing, and (2) the expectation of the appropriate case-marking in the context. Applying this line of reasoning to the oblique for nominative substitutions we can propose that (1) once other cues assign subjecthood to a nominal, there is a strong expectation of nominative case marking and (2) oblique cases have a relatively strong tendency to trigger case-related processing on their own. The implications for the oblique-for-oblique substitutions are

immediate: they should be intermediate in detectability between nominative-for-oblique substitutions and oblique-for-nominative substitutions. Th reason for this would be that they do not strongly violate case expectations, but they do involve relatively strong triggering of case related processing, leading to a cue clash. The relationship of the two hypothesized factors to error detectability is summarized in Table 7.

Error type:	Tends to violate case expectations that result from other cues?	Tends to triggers inappropriate case-related processing?
Nominative-for-oblique (the weakest violations)	**No** (suggesting that oblique contexts do not create strong expectations)	**No** (suggesting that nominative case does not strongly tend to trigger processing on its own)
Oblique-for-nominative (the strongest violations)	**Yes** (Since nominative case is not a strong processing trigger—see above—we conclude that nominative case tends to be expected when other cues indicate that a noun is subject.)	**Yes** (If oblique cases are relatively strongly prone to trigger processing on their own, then this will increase the strength of these violations.)
Oblique-for-oblique (intermediate in strength of violations)	**No** (since oblique contexts, as established above, do not create strong expectations)	**Yes** (since oblique cases were established above as relatively strongly prone to trigger processing independently)

Table 7. Two hypothetical factors to account for differing levels of sensitivity to different types of case errors.

It needs to be clear that in speaking of expectations we do not mean before-the-fact anticipations, but only that once semantic and/or discourse roles are determined for a subject noun, the processor wants the case form to converge with the other cues.

3.1.3.11. Processing considerations and aspect errors

Creating aspectual errors proved to be a bit of a challenge, in that often merely altering a form from perfective to imperfective or vice versa would not create a clear error. The analogous change in English might convert *He took off his shirt* to *He was taking off his shirt*, moving the event from foreground to background, but not producing clashing failure of cues to converge. As a matter of fact, wrong use of aspect *is* a notorious feature of nonnative Russian, at least anecdotally, but this may often involve less elementary functions of aspect. The two texts in the experiment did nevertheless provide some opportunities for altering aspect in ways that affected presumably basic functions and in the process creating sentences which native speakers uniformly reacted to as erroneous (although my primary consultant, had doubts about how native listeners would react until the data were in.)

Whether one relates aspect to inherent temporal properties of lexical items, or to foregrounding and backgrounding in discourse, the aspect substitutions in the experiment would be expected to cause problems. In the case of the perfective-for-imperfective

substitution, the verb was inherently stative (*stojat'*, 'be standing') and at best a poor candidate for perfective aspect, if perfective aspect is strongly associated with telic verbs. Likewise, in terms of discourse function, the clause was describing the background situation (that various items of furniture *stood* in the room) which simply could not be made into a foregroundable event. In the case of the imperfective-for-perfective substitution, the verb involved, *zakryt'*, 'shut' is an achievement verb and particularly so when combined with the particular direct object *glaza*, 'eyes'. Shutting ones eyes is a relatively punctual event. From the discourse perspective, this event was closely conjoined with another punctual event. The English analogue of this substitution would yield *He was closing his eyes and fell asleep* (in place of *He closed his eyes and fell asleep.*).

Participants in this experiment displayed extremely low sensitivity to the perfective-for-imperfective substitution (detections by only two participants in the Meaning-Oriented task), perhaps indicating a strategy that often simply ignores perfectivizing prefixes during lexical lookup. The imperfective-for-perfective substitution faired considerably better, but there was some doubt cast on five of the nine responses to this item in the Meaning-Oriented task. In the Form-Oriented task, roughly half of participants detected the imperfective-for-perfective substitution and roughly one third detected the perfective-for-imperfective substitution. This may indicate again that under relatively undemanding processing conditions there is more of a tendency for aspect to be processed than under more normal listening conditions such as those simulated in the Meaning-Oriented task. Instances when aspect is not implemented in comprehension would be parallel in principle to the assignment of semantic roles to nouns in the absence of the implementation of case inflection. For example, propositions in discourse would be assigned foreground and background status on the basis of considerations other than their aspectual marking.

3.1.3.12. Further (qualitative) consideration of the specific errors

Since there was only one instance of each specific error type and only two of each broader type (based on our theoretically motivated supracategories), we cannot confidently generalize to all errors of a given type. For that, we would have needed several instances of each specific type of error and this proved impractical. For now, the value of the present findings can perhaps be enhanced by a closer qualitative examination of the individual items. This is also important given the exploratory character of this research.

Items (3) and (4) were the items in which incorrect oblique case forms were substituted for the correct nominative forms (**oblique-for-nominative** substitutions):

(3)	Potom	*chelovek-om	stal	myt'	ruki	pod	kranom.
	then	person-instr	started	to.wash	hands	under	tap:instr

Then the man started to wash his hands under the tap.

Base-form: chelovek Correct target: chelovek Substituted: chelovekom

(4)	Teper'	ego	rubashka	i	*br'uk-ax	vis'at	na
	then	his	shirt:nom	and	pants-pl:loc	hang:pres:perf	on

stule,	a	sam	on	v	pizhame.
chair	and	self	he	in	pajamas

Then his shirt and pants are hanging on the chair and he himself is in his pajamas.

Base-form: br'uki Correct target: br'uki Substituted: br'ukax

Item (3) was discussed above as item (1). In (4) the overall complexity of the context in which the error occurs is worth noting. The incorrect form occurs as one conjunct of a conjoined noun phrase. For the native speaker this context might increase the impact of the error, if the case property belongs to the conjoined noun phrase as a whole, since the nominative-marked conjunct has already been encountered. The collocation *i br'ukax*, 'and pants:locative:plural' is quite possible, but would require a larger context such as *v rubask-e i br'ukax*, 'in a shirt-locative and pants-locative:plural' (which would be a common way of saying that someone was wearing a shirt and pants). We would expect that on encountering *br'ukax* in (4) the native mental language processor might attempt a reanalysis, but other cues (e.g., semantics, word order) would rapidly converge on a solution that was incompatible with any use of the form *br'ukax*. In other words, for native speakers, this should be a "good" error indeed (and my observation has been that native speakers react strongly to it). No Intermediate participants noticed it in the Meaning-Oriented task. In the Form-Oriented task, it was in fact the second most frequently noticed error for the Intermediates, behaving similarly to the other oblique-for-nominative substitution. In other words, although the two oblique-for-nominative substitutions involved substantially different syntactic contexts and substantially distinct sub-parts of the overall case inflection paradigm (plural vs. singular), they were comparable to one another in their tendency to be noticed by members of the Intermediate group in the less demanding of the two listening tasks. In the Meaning-Oriented task, the Advanced group detected the error in (3) more often than the one in (4), though not significantly so (thirteen detections and nine detections respectively). Thus in spite of the greater complexity of the error context of (4) in contrast with (3) and other important differences, these errors appear to be similar in their tendency to be detected under the conditions of the Meaning-Oriented task. In the Form-Oriented task the difference between the two errors for the Advanced participants was slightly larger (twenty-one detections in the case of item (3) and sixteen in the case of item (4). We cannot rule out the possibility that this difference is due to the difference in the complexity of the syntactic environments.

The **oblique-for-oblique** substitutions are presented here as (5) and (6).

(5)	Posle	togo	kak	namylil	ruki	*myl-e	on	polozhil	ego
	after	that	how	he.soaped	hands	soap:loc	he	put	it

v myl'nitsu.
into soap.dish
After he soaped his hands with the soap, he put it into the soap dish.

Base-form: mylo Correct target: mylom Substituted: myle

(6)	Kogda	on	lezhal	v	*krovat-ju,	on	popravil	podushku	pod
	when	he	lay	in	bed-instr	he	adjusted	the.pillow	under

golovoj.
head
As he lay in the bed, he arranged the pillow under his head.

Base-form: krovat' Correct target: krovati Substituted: krovatju

The incorrect form in (5), like all of the errors, was pronounced clearly on the tape. For a native speaker it differs from the base-form in the quality of the final vowel ([ɪ] rather than [ə] and of the consonant proceeding the final vowel ([ʲ] rather than [l], which is in turn detectable in the preceding vowel—see section 1.3.1 above). For nonnative speakers, both of these contrasts are potentially difficult to discriminate, depending on the L1. However, it should be relatively easy to discriminate the incorrect, locative form from the correct target instrumental form (final syllable [ʲɪ] rather than [ləm]). For a language processor that is making use of case inflection, *myle*, if simply equated perceptually with the base-form *mylo*, would be still be unable to find a compatible semantic role. If the deviation from the base-form were not perceived, this would in essence amount to another nominative-for-oblique substitution, rather than oblique-for-oblique.

For L1 English participants, the nominative/accusative ending and locative ending are, impressionistically at least, reasonably discriminable on the basis of the vowel, though not necessarily on the basis of the consonant. If the reason for the low number of detections of this error in the Meaning-Oriented task (three detections by the Advanced group, none by the Intermediate group) were primarily perceptual, then we would not expect a dramatic increase in the Form-Oriented task. In fact, we find the number of detections by the Advanced group increasing substantially in the Form-Oriented task (increasing from one detection to seven). Thus while perceptibility may be a factor, it is does not appear to be an excessively powerful factor. In the phonetically parallel case of the nominative-for-locative substitution, there was a smaller increase going from the Meaning-Oriented to Form-Oriented task (from two to five for the Advanced group), as predicted by our processing considerations. In other words, although perceptibility is a concern in these cases, the performance of the participants is more consistent with a deeper processing explanation than with a perceptual explanation.

The low performance with (5) in the Meaning-Oriented task may be an artifact of the experiment. This was the only text segment containing an error that immediately followed another text segment containing an error. Participants may thus have been presented with this second error during a refractory period when they had not yet recovered from reacting to (or equivocating over) the previous error. The previous error (the first oblique-for-nominative substitution, *chelovek-om*, 'person-instrumental' for *chelovek*, 'person:nominative') was in fact the case marking error to which there was the highest sensitivity overall. If reactions to errors did create a refractory period, then (5) occurred in a particularly infelicitous location. The fact that the tendency to notice this error increased so much in the Form-Oriented task, where the interstimulus interval was nearly twice as long, supports this explanation.

Item (6) in the Meaning-Oriented task is the datum that appears to be behaving the most inconsistently with expectations. As an oblique-for-oblique substitution, the prediction was that the number of detections would be intermediate between the higher number of detections of oblique-for-nominative substitutions and the lower number for

nominative-for-oblique substitutions. Yet it was the most frequently detected case marking error in the Intermediate group, (two detections) and the Advanced group (six detections) during the Meaning-Oriented task. This could have various explanations. Within the confines of the data available to us, we can apparently rule out the fact that the noun is governed by a preposition. The erroneous word in (8) below—*v kresl-o*, 'in chair-nominative'—is also governed by a preposition, in fact the same preposition. It is true that *v kresl-o*, 'in chair-nominative' is a possible collocation (due to the neutralization of the nominative/accusative distinction in neuter nouns; the meaning would then be 'into the chair'), while the error in (6) is not a possible collocation. However, the full conjoined noun phrase *rubashka u br'ukax*, 'shirt:nominative and pants:locaive:plural' in (4) also involves an impossible collocation, and yet its level of detectability conforms to our expectations. It is possible that the combination of prepositional government and an impossible collocation is responsible. Perhaps a better possible explanation is that the specific correct target with its preposition, *v krovati*, 'in bed' is a familiar phrase. Familiar phrases may create the expectations of the specific word form. That would mean that the error in (6) happens to involve both an expectation of a particular form in the context and an oblique case as the erroneous form, triggering case-related processing. Along the lines of reasoning embodied in Table 7, this would make the error in (6) comparable to an oblique-for-nominative substitution, predicting relatively high detectability.

Finally, detectability of this error may be an artifact of the experiment, in which the noun *krovat'*, 'bed' happens to occur more frequently than the nouns involved in the other errors. The lexeme had already occurred in the text, both in the base-form (once) and the more general oblique form *krovati*, 'bed:dative/locative/genitive' (five times). The other nouns containing errors had each occurred only once previously to the error, with the exception of *br'ukax*, 'pants', which had not occurred previously. I had some concern that the lack of repetition would seriously disadvantage *br'ukax*, but this appears not to have been the case. However, the situation with *krovat'/krovati/krovatju* may be rather extreme in terms of the amount of repetition.

Any of these explanations would count as an argument against a decisive role for phonetic salience since, as noted, the relative advantage of phonetic salience over case function as a predictor of the rank order of error detections in the Meaning-Oriented task was in large part due to this one item. It may also be important that in the Form-Oriented task, with its increased time for processing and without the divided attention feature of the Meaning-Oriented task, item (6) takes its predicted place in the rank ordering predicted on the basis of case processing.

Turning to the **nominative-for-oblique** substitutions:

(7) Nakonets chelovek vyter ruki *polotentse,
finally person wiped hands towel

i voda perestala kapat'.
and water stopped to.drip
Finally, the man wiped his hands with a towel and the water stopped dripping.

Base-form: polotentse Correct target: polotentsem Substituted: polotentse

(8) Chelovek ochen' ustal sidet' v *kreslo pered televizorom.
person very tired to.sit in armchair before television.
The person grew very tired of sitting in the armchair in front of the television.

Base-form: kreslo Correct target: kresle Substituted: kreslo

In both (7) and (8) the listener encounters the base-form (nominative/accusative), clearly enunciated, in an impossible context. The possible difficulty in discriminating *kreslo*, 'chair:nominative/accusative' and *kresle*, 'chair:locative' is parallel to that discussed in connection with item (5), where the dramatic improvement in the Form-Oriented task suggested that perceptibility was not a major issue. I suspect that the best explanation for the low number of detections of this error is that nonnative speakers are again simply not implementing the inflectional form. And in the case of (7), perceptibility would not appear to be an issue (*polotentse*, 'towel:nominative/accusative' vs. *polotentsem*, 'towel:instrumental'). This is the error that two native controls, one of whom was observably distracted, failed to detect. The context is a reasonable one for a top-down mental restoration (Samuel, 1981; Marslen-Wilson, 1973) of the correct form by native listeners: clause final position, highly constraining semantics and phonological reduction. For the experimental participants, in view of the overall findings of the experiment, the possibility of a top-down mental restoration of the correct ending would seem to be out of the question. Rather, it appears that they tended not to implement the inflection.

Finally, there were the two items involving substitutions of the **incorrect verbal aspect** in place of the correct one:

(9) I poshol v spal'n'u, gde *po-stojali krovat', stolik, i stul.
and he.went into bedroom where perf-stood bed table and chair
And he went into the bedroom where there stood a bed, a table and a chair.

Correct target: stojali Substituted: postojali

(10) Potom on *zakr-yva-l glaza i usnul
then he close-impf-past eyes and went.to.sleep
Then he closed his eyes and went to sleep.

Correct target: zakryl Substituted: zakryval

The expectation here was that there would be more of a tendency among Intermediate participants to notice the replacement of an imperfective form by a perfective form than vice versa, based on a strong tendency I observed in the spoken production of nonnative speakers at a comparable level to overuse imperfective forms in spoken narrative production. On the same basis, I expected an increased tendency in the Advanced group to react to substitutions of imperfective for perfective. On the other hand, in terms of the functional considerations discussed earlier, both of these errors clash both with the inherent lexical aspect of the verbs and with the discourse functions (foregrounding and backgrounding). As discussed above in connection with the topic of markedness, the situation here is more complex than with noun inflection, and both errors may have essentially substituted a marked erroneous form for an unmarked contextually required form. In any case, only a single participant in the Intermediate group detected only a single error in the combined items (9) and (10) during the Meaning-Oriented task. The Advanced participants did show a difference in reactions to (9) and (10) during the

Meaning-Oriented task, in the opposite direction from that expected: two detections (ten percent of participants) in the case of (9) and eight detections (forty percent of participants) in the case of (10). Keep in mind, however, that the number of detections may have been somewhat inflated due to a subject indicating an identical error aloud in a group of seven participants.

Item (9) is particularly fascinating. The imperfective verb *stojat'*, 'stand (imperfective)', is an extremely high frequency lexical item (ranked 104 in Brown's [1996] frequency ranking of 10,000 high frequency Russian words). The prefix *po-* is also an extremely common mark of perfectivization. Even the erroneous item *postojat'*, 'stand for awhile' (perfective) is relatively frequent (ranked 1,653 in Brown, 1996). In fact, fifty-nine percent of participants detected and corrected this error in the Printed sentences task. Nevertheless, in the Meaning-Oriented task, only five percent of participants (two out of thirty-seven) detected it. Regardless of the exact status of *postojat'* in relationship to *stojat'* (i.e., whether or not they are simple aspectual partners), these results certainly hint at a lexical access strategy in which the prefix is ignored and the lexeme accessed via the stem, reminiscent of the affix stripping hypothesis for English visual word processing (Taft, 1979; Taft & Forster, 1976; Lima, 1987), except without further utilization of the particular prefix. Impressionistically, it would seem that the auditory difference *postojat'* and *stojat'* would be very difficult to miss. The lack of sensitivity to item (9) on the part of participants, then, would appear to strongly support a claim that the failure to implement inflectional form can occur even when inflectional form is clearly perceptible.

3.2. EXPERIMENT 2

In the most general terms, the results of Experiment 1 are compatible with the claim that nativelike inflectional processing in adult learners of Russian as a second language develops, but develops slowly. While there is little reason so far to believe that the system would ever develop to a nativelike or near-nativelike level. On the other hand, development does appear to be taking place. In particular, in the Meaning-Oriented task, where there is little scope for the application of metalinguistic knowledge to the error judgments, there is a significant difference between the Intermediate and Advanced groups. This is also true in the so-called Form-Oriented task, where I have argued that performance still depends on normal language processes to a significant extent. The present experiment will attempt to replicate those findings.

In addition, the pattern of reactions to errors in Experiment 1 appeared to be systematic, within the limits of the statistical tests used (Spearman ρ and χ^2) and the specific errors examined. Moreover, the systematicity conforms to theoretically motivated predictions, allowing at least some grounds for optimism that the results are pointing to something that is both valid and interesting. Experiment 1 dealt with nominative case and two oblique cases. Part of the reasoning behind that choice was that the nominative-oblique contrast seemed in some sense a more extreme contrast than nominative-accusative or nominative-dative contrasts. For the first experiment it was felt that using the presumed stronger contrast would increase the chances of detecting processing differences. Experiment 2 first of all attempts to replicate the findings of Experiment 1 in relation to the differences between group means and task means. More centrally, Experiment 2 broadens the picture to include the nominative-accusative contrast, verbal person agreement and verbal gender agreement. Thus in Experiment 2 the focus is more on the *pattern* in the developing sensitivity to inflection than the *existence* of increasing sensitivity. That is, we are concerned here especially with the third hypothesis stated in the

early part of this chapter: *There will be systematic patterning in the tendency of participants to notice some error types earlier or more often than others*, and this tendency will be related to functional considerations of the sort discussed. We hope to extend our picture of the systematic pattern that may have begun to emerge in the results from Experiment 1.

In addition to the new error categories employed in Experiment 2, three errors were repeated from Experiment 1.

3.2.1. Method

3.2.1.1. Participants

Fifty-four nonnative speakers of Russian residing in St. Petersburg, Russia, participated in the experiment, as well as twelve native controls. The grouping criterion was applied from the outset: those who began learning Russian less than four years previously were placed in the Intermediate group and those who had been learning Russian four years or more were placed into the Advanced group. The mean period of time since the onset of Russian learning for the Advanced group was 7.07 years (ranging from four years to seventeen years, though in the latter case there was a 12 years hiatus). The mean time since onset of Russian learning for the Intermediate group was 2.02 years (ranging from six months to 3.7 years, with only one participant having started less than one year earlier). In this experiment, there turned out to be no relationship between time since onset of learning and scores in the Meaning-Oriented task ($r = .17$, $p > .05$). There was a small but significant relationship between years since onset of Russian learning and scores in the Form-Oriented task ($r = .28$, $p < .05$). On the other hand, as in Experiment 1, there was clearly no relationship between years in Russia and performance on either the Meaning-Oriented task ($r = .09$) or the Form-Oriented task ($r = .02$). Thus once again, it appears that grouping in terms of time since onset of Russian learning was a better criterion than years in Russia, although also once again, this probably resulted in some misassignments. Yet as the overall results indicate that the grouping criterion was adequately valid for detecting developmental differences. The participants had lived in Russia an average of 1.7 years (Advanced participants) and 1.38 (Intermediate participants). Only three participants (two Intermediate, one Advanced) had lived in Russia less than six months.

The L1s of the participants included Chinese: 13 (2 Intermediate, 11 Advanced); Korean: 9 (5 Intermediate, 4 Advanced); Dravidian languages: 6 (Intermediate); Arabic: 6 (Intermediate); English: 5 (3 Intermediate, 2 Advanced); Japanese: 4 (3 Intermediate, 1 Advanced); Indo-Aryan languages: 3 (Intermediate); Icelandic: 1 (Intermediate); Spanish: 1 (Advanced); French: 1 (Advanced); Danish: 1 (Advanced); Finnish: 1 (Intermediate); and Farsi: 1 (Intermediate); Italian: 1 (Advanced); and Thai: 1 (Intermediate). All participants were university students, representing age-groups twenty years or under: 8; twenty-one to thirty: 38; thirty-one to forty: 5; over forty: 2 (one did not report her age group). None of the participants began learning Russian before adolescence. Four appear to have begun learning Russian in high-school. There were twenty-six female participants and twenty-eight male participants.

In addition, as noted, there were twelve native Russian control participants. Eight were students in a private English language program, one was a native Russian speaker from abroad who had come to Russia to acquire Russian literacy, and three were students in post-secondary institutions.

3.2.1.2. Materials

The same two basic texts from the first experiment were modified for the second experiment and re-recorded. The pictures used for the Meaning-Oriented task were rearranged so as to guarantee that turning pages would not interfere with detecting errors. The same female speaker recorded the texts.

Three of the errors employed in this experiment were taken from Experiment 1. They were included in an effort to correct possible problems with the first experiment, and also to provide a possible point of comparison between the experiments. The two oblique-for-oblique substitutions were repeated. Item (5) of Text 1 (locative-for-instrumental substitutions) was presented without there being an error in the immediately preceding segment, in order to rule out the possibility of a refractory period interfering with responses. Segment (14) of Text 2 (item 10 above) was repeated and the number of previous tokens of the word *krovat'* 'bed' was reduced by substituting the synonym *postel'* for it in two places, in view of the possibility that in Experiment 1 the high token frequency of *krovat'* had increased sensitivity to the lexeme. Thirdly, the second imperfective for perfective substitution, segment (16) in Text 2, was included again, due to the fact that a participant had reacted out loud to the same error in segment (7) of Text 1, potentially influencing six other participants who also noted the error.

Six errors were new to Experiment 2: two accusative-for-nominative substitutions, two nominative-for-accusative substitutions, one error of person agreement (erroneous second person singular replacing correct third person singular) and one error of verbal gender agreement (erroneous masculine replacing correct target feminine). An error was also introduced into the first segment of Text 1 so that it could be used for training participants in the procedure. Thus, there were a total of ten errors in twenty-seven sentences, nine of which were tabulated, with never more than one error per segment, no errors in initial or final words of segments and no two segments in a row containing errors. Answer sheets with instructions were provided for each task. For the third task, the answer sheet contained, in printed form, those sentences from the texts that contained errors. The texts used for Experiment 2, with errors in boldface and correct forms in parentheses, are included in the Appendix.

3.2.1.3. Procedure

The procedure in Experiment 2 was the same as in Experiment 1.

3.2.2. Results

Native Russian speaking controls performed only the Meaning-Oriented task. As a group, they noticed ninety-nine percent of the errors. Only one error was missed, by a single participant (the imperfective-for-perfective substitution, Text 2, segment 16). The two items that received unexpected error judgements by control participants in Experiment 1 were modified in Experiment 2. In Experiment 2, no sentences were judged by native controls to contain errors apart from those in which errors were present by design. Thus with one exception involving one item, control participants detected all errors and did not react to any sentences that did not contain errors. The mean score of native control participants in the concurrent picture-selection task was 99.3%

Turning to the experimental groups, in the Meaning-Oriented and Form-Oriented tasks, reactions to errors were only counted if the Printed Sentences task demonstrated that the participant was able to identify the precise error involved. This reduced the number of possible cases where participants might have marked an "X" correctly as a

result of guessing. This affected 14.7% of the correct responses in the Meaning-Oriented task, and 14.2% in the Form-Oriented task. This reflects the higher tendency to guess in Experiment 2 than in Experiment 1.

Looking at the two groups and three tasks as a whole, an analysis of variance revealed significant main effects of both task ($F[2,52] = 159.601$, $MSe = 2.374$, $p < .001$) and level ($F[1,52] = 15.110$, $MSe = 9.229$, $p < .01$) and a significant interaction of task by level ($F[2,104] = 3.187$, $MSe = 2.374$, $p < .05$) as summarized in Table 8.

Source	df	Sum of Squa...	Mean Square	F-Value	P-Value
Group	1	139.451	139.451	15.110	.0003
Subject	52	479.913	9.229		
Task	2	757.690	378.845	159.601	.0001
Task * Group	2	17.344	8.672	3.653	.0293
Task * Subject	104	246.866	2.374		

Dependent: Task

Table 8. Experiment 2: ANOVA source table, three experimental tasks performed by two groups (Intermediate and Advanced L2 Russian users, as defined), with participants as the repeated measure.

A separate analysis was done in which experimental items rather than participants were used as the repeated measure (percentage of participants detecting each error was used as the dependent measure, rather than absolute tallies, due to uneven group sizes). Again the effects of task ($F[2,24] = 64.343$, $p < .001$) and level ($F[1,24] = 40.709$, $MSe = 359.852$, $p < .001$) were significant. The interaction of task by level approached significance ($F[2,24] = 3.085$, $MSe = 64.343$, $p = .0642$).

Figure 2 summarizes the findings graphically. We turn next to specific comparisons.

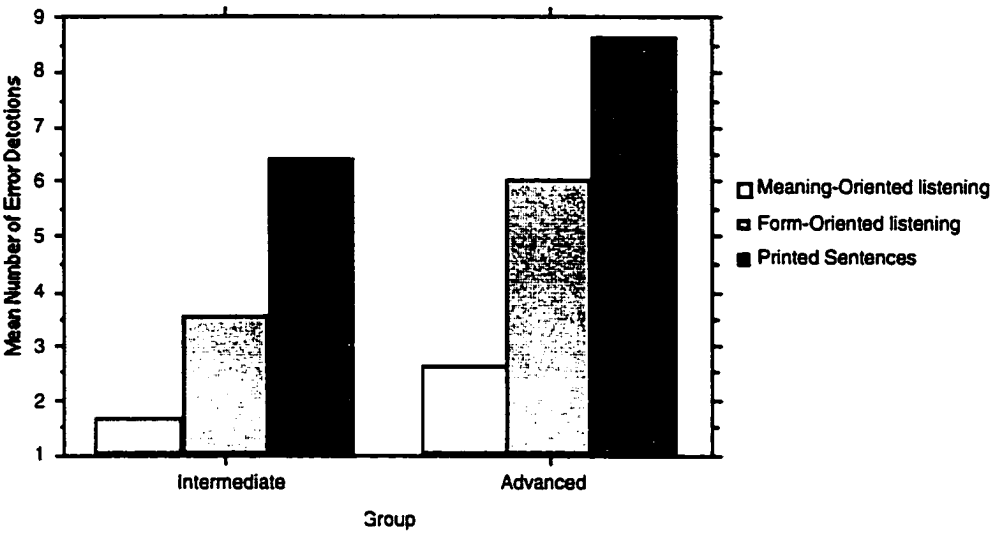

Figure 2. Experiment 2: Error detections in three tasks by the two groups. (The maximum number of errors detected was nine.)

The mean correct responses for experimental participants in the picture selection subtask of the Meaning-Oriented listening task was 89.4%, the Intermediate mean being somewhat lower than the Advanced mean (87.7% and 93.3% respectively, $t[53] = 1.986$, $p = .0523$, 2-tail).

3.2.2.1. Meaning-Oriented Listening

Participants noticed an average of 2.02 errors (22.4%) out of the nine possible. For the Intermediate group the mean error detections was 1.63 (18.1%) out of the possible nine (*St. Dev.* = .370, range from 0 to 7). For the Advanced group the mean was 2.59 (28.8%, *St. Dev.* = 2.364, range from 0 to 8). Six of the twenty-two Advanced participants noticed more fifty percent or more of the nine errors, as did three of the thirty-two Intermediate participants. Of the Advanced participants, 63.6% detected one third or less of the errors, as did 78.1% of the Intermediate participants. The difference in mean error detections between the Intermediate and Advanced groups (1.63 and 2.59, respectively) in the Meaning-Oriented task, however, was not significant ($t[52] = 1.5817$, $p = .1198$, 2-tail).

3.2.2.2. Form-Oriented Listening

In the second task, participants detected an average of 4.56 errors (50.6%) of the nine possible ($SD = 2.6823$, range from 0 to 9), an increase of 2.54 errors (28.2%). The mean detections for the Intermediate group was 3.53 (39.2%) out of nine ($SD = 2.7355$, range from 0 to 9), an increase of 1.91(21.2%). For the Advanced group, the mean errors detected was 6.0455 (67.2%) out of nine ($SD = 1.7856$, range from 3 to 9), an increase of 3.45 errors (38.4%). The difference between the Intermediate group and the Advanced group (means 3.53 and 6.05 respectively) was significant ($t[52] = 3.7863$, $p < .001$).

3.2.2.3. Incorrect error judgments (false alarms)

In the Meaning-Oriented task, there was an average of 1.76 incorrect error judgements per participant. This amounted to a nearly one-to-one ratio of correct error detections to incorrect guesses. In the Form-Oriented task there were an average of 4.06 false alarms per subject. This ratio of correct error detections to incorrect cases was somewhat less than one-to-one (219 false alarms to 246 error detections). For the reason discussed earlier, this tendency to attribute errors to sentences independently of the presence of errors weakens our ability to discern differences in sensitivity to different errors and error types (note that the tendency to react when errors were present as approximately double the tendency to react when no error was present). At the same time, the increase in false alarms from the first task to the second task does suggested that there was a greater focus on form in the second task than in the first, as intended.

3.2.2.4. Printed Sentences

In the Printed Sentences task, participants reacted to an average of 7.30 (81.1%) errors per subject (range from 0 to 9, $SD = 2.116$) out of the possible nine. The mean number of error detections in the Intermediate group was 6.41 (71.2%; range from 0 to 9, $SD = 2.213$). For the Advanced group the mean was 8.60 (95.6%; range from 7 to 9; $SD = .666$). The difference between the two groups in the Printed Sentences task was significant ($t[52] = 2.275$, $p = .027$). Twenty of the twenty-two Advanced participants discovered eight or nine of the errors, as did twelve of the thirty-two Intermediate participants. Of the Advanced participants, none reacted to less than seven of the nine text segments containing errors, while fourteen of the thirty-two Intermediate participants reacted to less than 7 (5 reacted to less than half).

3.2.2.5. Printed Sentences Task and Listening Tasks

The Printed Sentences task was intended to measure of off-line, metalinguistic knowledge. If scores in either of the other two tasks were to be unrelated to scores in the Printed Sentences task, this would suggest an inability to apply metalinguistic knowledge in that task. However, even in the case of the Intermediate group in the Meaning-Oriented task, there was a relationship of scores in that task to scores in the Printed Sentences task ($r = .479$, $p < .01$). This does not necessarily mean that the participants were applying metalinguistic knowledge in the Meaning-Oriented task, but raises the possibility. By contrast, the relationship between the Printed Sentences task and the Meaning-Oriented task for the Advanced group was small and nonsignificant ($r = .191$, $p > .05$). The relationship between the Printed Sentences task and the Form-Oriented listening task was significant for the Intermediate group ($r = .653$, $p < .01$) but not for the Advanced group ($r = .137$, $p > .05$). The lack of correlation between the Printed Sentences task and the other tasks in the case of the Advanced group can be attributed to the fact that there was a ceiling effect with that group in the Printed Sentences task (mean = 95.6%; range from 7 to 9; SD = .666).

3.2.2.6. Error types and error detection

The instances of error detection were not randomly distributed among the errors, nor among the error types. Tables 9 and 10 summarize these findings for the three experimental tasks.

	Meaning-oriented Task		Form-oriented Task		Printed Sentences task	
Error Type:	Inter-mediate group	Advanced group	Inter-mediate group	Advanced group	Inter-mediate group	Advanced group
Nominative (chelovek) replaced by accusative (cheloveka)	7 (21.9%)	9 (40.9%)	16 (50%)	12 (54.5%)	27 (84.4%)	22 (100%)
Nominative (muzhchina) replaced by accusative (muzhchiny)	7 (21.9%)	8 (36.4%)	14 (43.6%)	17 (77.3%)	28 (87.5%)	21 (95.5%)
Locative (krovati) replaced by instrumental (krovatju)	7 (21.9%)	10 (45.5%)	14 (43.6%)	18 (81.8%)	20 (62.5%)	22 (100%)
Instrumental (mylom) replaced by locative (myle)	5 (15.6%)	6 (27.3%)	12 (37.5%)	11 (50%)	24 (75%)	21 (95.5%)
Accusative (lampu) replaced by nominative (lampa)	4 (12.5%)	8 (36.4%)	16 (50%)	18 (81.8%)	23 (71.9%)	21 (95.5%)
Accusative (rubashky) replaced by nominative (rubashka)	8 (25%)	5 (22.7%)	16 (50%)	15 (68.2%)	28 (87.5%)	21 (95.5%)
Third person (sidit) replaced by second person (sidish)	11 (34.4%)	12 (54.5%)	19 (59.4%)	20 (90.9%)	27 (84.4%)	22 (100%)
Perfective (zakryl) replaced by imperfective (zakryval)	2 (6.3%)	3 (13.6%)	3 (9.4%)	9 (40.9%)	9 (28.1%)	16 (72.7%)
Feminine verb agreement (perestala) replaced by masculine (perestal)	3 (9.4%)	3 (13.6%)	8 (25%)	12 (54.5%)	17 (53.1%)	21 (95.5%)

Table 9. Errors detected in Experiment 2. In each Intermediate group cell the maximum number of possible detections was always thirty-two, and in each Advanced group cell—twenty-two.

The distributions of scores among error types appears to be non-random ($\chi^2[8] = 93.397$, $p < .001$ for the Meaning-Oriented task and $\chi^2[8] = 72.025$, $p < .001$ for the Form-Oriented task). The nonrandomness of the distribution appears to be mainly due to the verbal inflectional errors, one of which is the most detected error overall, and the other two of which are the least detected errors. Even if the verb inflections are examined separately, the distribution of scores appears to be non-random (for the Meaning-Oriented task $\chi^2[2]= 16.546$, $p = .0003$ and for the Form-Oriented task, $\chi^2[2]= 16.2512$, $p = .0003$). On the other hand, if the noun inflections are examined separately, the distribution appears to be random ($\chi^2 (5)= 2.374$, $p = .759$ for Meaning-Oriented listening and $\chi^2 (5)= 3$, $p = .7$ for Form-Oriented listening).

Relating the performance of the two groups of participants on the various error types, one finds a rank order correlation between the two groups (Spearman ρ, corrected for ties) of .687 for the Meaning-Oriented listening task, which approaches significance

(p = .051) and .734 for the Form-Oriented Listening task, which is significant (p = .038). However, if the comparison is limited to the noun inflectional errors, there is no rank order correlation between scores of the two groups (ρ = .123, p = .783 and ρ = .235, p = .50 for the Meaning-Oriented task and the Form-Oriented task, respectively). Here too, there appears to be a systematic patterning of the verbal inflectional errors, but not of the nominal inflectional errors.

Type of Substitution	Error detections in Meaning-Oriented Task	Error detections in Form-Oriented Task	New error detections in Form-Oriented task
Accusative replacing nominative	31 (28.7%)	59 (54.6%)	28 (25.9%)
oblique replacing oblique	28 (25.9%)	55 (50.9%)	27 (25%)
Nominative replacing accusative	25 (23.2%)	65 (60.2%)	40 (37%)
χ^2 value	$\chi^2(2) = .643$, p = .725	$\chi^2(2) = .849$, p = .654	$\chi^2(2) = 3.305$, p = .197

Table 10. Experiment 2: detections of case errors by general type of substitution, combining the Intermediate and Advanced groups.

Table 10 summarizes the differences in case error types. These finding make it impossible to reject the null hypothesis of no differences in sensitivity to different error types. There may be a trend toward a greater tendency to detect errors in going from the Meaning-Oriented task to the Form-Oriented task with respect to the nominative-for-accusative substitutions (forty detections vs. twenty-eight and twenty-seven for the other error types), but it is non-significant.

3.2.2.7. Phonetic salience and error detection

In Experiment 2 it appears less likely that phonetic salience is playing a role. The two accusative-for-nominative substitutions are extremely similar in their tendency to be detected, although they presumably differ in phonetic salience (the addition of a syllable vs. a vowel change). The gender agreement error could be argued to involve at least as salient an inflectional form as the person agreement error (the addition of a syllable vs. a consonant change), and yet they strongly differ in the direction opposite from what one might expect if phonetic salience were at the root of the difference in error detectability.

3.2.3. Discussion

In Experiment 2, we again see a striking difference between native controls and nonnative participants in their levels of sensitivity to inflectional errors, and by implication, a difference in their tendencies to process inflection. We also once again see evidence that nonnative speakers become more sensitive to inflectional errors over time, which opens the possibility that development in the direction of nativelike processing is going on. The task-by-group interaction probably reflects the fact that Advanced participants benefited more than intermediate participants from the decreased processing demands of the Form-Oriented task vis-à-vis the Meaning-Oriented task. Whereas Intermediate participants increased an average of 1.90 error detections going from the first task to the second task, Advanced participants increased an average of 3.45 errors.

We argued in connection with Experiment 1 that much of the improved performance in the Form-Oriented task (in comparison with the Meaning-Oriented task) should be attributed to the decreased processing demands of the second task: the text is already familiar; there is no secondary picture selection task; and the interstimulus interval is nearly twice as long. The task-by-group interaction seen in Experiment 2 can be taken as further evidence for the role of normal linguistic (as opposed to metalinguistic, analystical) processes in the improvement seen in the Form-Oriented task, on the assumption that the tendency of learners to process inflection increases over time. That is, easier processing conditions cannot benefit participants beyond what their underlying inflectional processing ability will allow. As learners come to have more underlying inflectional processing ability, its presence is uncovered by allowing more scope for it to operate, that is, by decreasing processing demands.

In Experiment 2, the only comparison that was not significant was the between-groups comparison of mean error detections in the Meaning-Oriented task. This can largely be attributed to participants from Korea, of whom there were nine altogether, divided almost evenly between the Intermediate and Advanced groups (4 and 5 participants respectively). Without the participants from Korea, the difference between group means in the Meaning-Oriented task approaches significance ($t[43] = 1.998$, $p = .052$, 2-tail). It is possible that the exceptionally low performance of the Advanced Korean participants is the result of cross-language influence. However, given the linguistic diversity of participants in general, this seems unlikely. Perhaps a better possibility is that years of Russian study in Korean universities did little to develop listening comprehension mechanisms.[20] The Korean participants' mean number of error detections in the Printed Sentences task was 7.56, compared with the mean of 7.3 for all participants. This suggests the Korean participants possessed much of the relevant metalinguistic knowledge and/or L2 reading strategies, but were less developed than non-Koreans, on average, in more fundamental language processing mechanisms. This explanation would need to be verified through an examination of Russian programs in Korea, but is consistent with reports from numerous Koreans regarding English language learning programs in Korea.

All in all, participants were able to identify and make a credible (in fact, almost always precisely correct) attempt at correcting 394 out of a possible 486 errors in the Printed Sentences task. In essence, that sets a limit on which errors could potentially have been detected in the other tasks. Out of 394 possible error detections, there were 109 actual error detections in the Meaning-Oriented task and 246 in the Form-Oriented task. It should be emphasized once again that the errors involved the most elementary uses of the inflections in question. In a sample of inflectional forms in contexts representing a wider range of uses, we would expect participants to detect a smaller percentage of errors than they did in the present experiment. Although these results are stronger than the results we saw in Experiment 1, they still indicate that the development of inflectional processing is a relatively long-term prospect, at best. The improved performance with the case errors in Experiment 2 vis-à-vis Experiment 1 could be due to the relatively high detectability of nominative-for-accusative and accusative-for-nominative substitutions (by the developmental stage represented by the Intermediate participants). It might also need to be argued that having more easily detectable noun errors overall somehow improved performance in the oblique-for-oblique substitutions in Experiment 2 in contrast with Experiment 1. However, as noted, the high incidence of guessing might be enough to

[20] This is an example where grouping subjects on the basis of an independent measure of proficiency would no doubt have yielded some different assignments from those that resulted from our use of *time since onset of learning* as a proxy for developmental levels.

cancel out the difference between relatively highly detectable (nominative-for-accusative; accusative-for-nominative) and moderately detectable errors (oblique-for-oblique, perhaps).

3.2.3.1. The pace of development

After several years of Russian learning, including on average over sixteen months of life in Russia, usually including participation in university courses taught in Russian, participants in the advanced group reacted to less than thirty percent of the errors during the Meaning-Oriented task. There was considerable individual variation in Experiment 2, with six of the twenty-two Advanced participants and even three of the thirty-two Intermediate participants detecting over fifty percent of the errors. In fact, one of the Advanced participants reacted to eight of the nine errors, just like one of the native controls. Two of the Intermediate participants detected seven errors in the Meaning-Focussed task. This slightly alters the picture from Experiment 1, opening the possibility that language processing mechanisms in a few participants may develop at a more rapid pace than normal. In general the picture remains the same, in that nonnative performance is far from native performance after several years of language learning. We also need to keep in mind the simple nature of the two texts and the elementary nature of the errors presented in the experiments.

3.2.3.2. Processing considerations and error detections

Experiment 2 again employed errors of noun-inflection and errors of verb-inflection. In the case of noun-inflection, a trend may be visible for nominative-for-accusative errors to be more detectable than the others (particularly seen in the new detections in the Form-Oriented task—see Table 10), but given the null hypothesis of an equal tendency for all error types to be detected, the distribution of error types did not significantly differ from chance, as we saw. The distribution of detections of verbal errors did differ significantly from chance.

3.2.3.2.1. Case errors

In the observed distribution in the Meaning-Oriented task, the accusative-for-nominative substitutions were noticed more often than the nominative-for-accusative substitutions (28.7% to 23.2%). This order reverses in the Form-Oriented task (54.6% to 60.2%). If there is in fact a difference in sensitivity to these error types, it is not robust enough to overcome the error added to the data by the high incidence of guessing in Experiment 2. However, the experiment was in fact sensitive enough to detect differences in responses to verb errors (and differences between verb errors and noun errors). Therefore we will tentatively assume that there is no difference between sensitivity to nominative-for-accusative errors vs. accusative-for-nominative errors. (The slight advantage observed for nominative-for-accusative substitutions in the Meaning-Oriented task, if genuine, could be due to the fact that some of the participants are in the habit of self-monitoring for—and self-correcting—nominative-for-accusative substitutions in their own spoken production.)

The situation with the oblique errors is complicated by the fact that our instrumental-for-locative substitution continued to trigger a relatively large number of reactions. Nevertheless, even if the instrumental-for-locative substitution is removed from the picture the distribution of error types still does not differ from chance ($\chi^2[2]= 3.890$, $p = .143$ for the Form-Focused task, in which 42.6% of participants reacted to the

instrumental-for-locative substitution, as opposed to 54.6% and 60.2% with the accusative-for-nominative and nominative-for-accusative substitutions, respectively).

In the general discussion below, these findings will be considered in combination with the findings in Experiment 1, and the hypotheses embodied in Table 7 and related discussion will be modified accordingly. Taken by themselves, these findings in Experiment 2 do not enable us to reject the null-hypothesis of equal likelihood of all case-error detections.

3.2.3.2.2. Verb errors

In the case of the verbal inflectional errors, there were significant differences between the person agreement error, on the one hand, and the gender agreement error and the aspectual error on the other. Sensitivity to the error of person agreement was relatively high. It was detected by 40.7% of participants in the Meaning-Oriented task and 72.2% in the Form-Oriented task (90.9% of Advanced participants, or twenty out of twenty-two, in the Form-Oriented task). Compare this with 9.3% for the aspect error and 11.1% for the gender agreement error in the Meaning-Oriented task, or 22.2% for the aspect error and 33% for the gender agreement error in the Form-Oriented task (40.9% and 54.5% rexpectively, for Advanced participants in the Form-Oriented task). We noted in the previous chapter that person agreement plays a role in both discourse level topic maintenance (or topic change) and at the level of individual frames in discourse models, identifying one argument of the verb. This is also true of person agreement. However, there are differences. Second-person agreement (used in the experiment) also consistently signals the place of a central speech act participant as one of verbal arguments. It lacks the degree of discourse-anaphoricity (and hence referential variability) of gender agreement. Thus it should not be surprising to find that the second person agreement inflection robustly triggers inflectional processing related to the identity of the subject (focal participant in the mental model), leading to a cue clash when the subject is third person singular and the verbal agreement is second person singular. In the terms developed in connection with sensitivity to case errors (see Table 7), the strong detectability of this error would be consistent with a simultaneous strong expectation of the correct inflection and strong tendency of the actually encountered inflection to trigger processing. I find that to be an intuitively plausible hypothesis for person agreement.

With gender agreement, the identity of the subject is constrained less strictly and less directly than with person agreement. That is, a particular gender ending, it least in the case of inanimate nouns, does not directly constrain the identity of the focal participant in the mental model, but rather constrains it via the word by means of which it was introduced. That is, a single entity can be referred to by nouns belonging to more than one gender (*kniga*, 'book' is feminine, while *slovar'*, 'dictionary' is masculine; *sobaka*, 'dog' is feminine, while *zhivotnoe*, 'animal' is neuter). Such an indirect form-to-concept relationship is inherently more complex than a simple deictic relationship (a relationship of linguistic form to a relatively stable aspect of non-linguistic experience) and hence the necessary mental associations should be more difficult to form in the case of gender agreement than in the case of person. It would be interesting to investigate whether there would be more sensitivity to gender violations involving semantic gender, especially when gender is tied to the sex of a human participant. In any case, we would ultimately want to account for the difference in sensitivity to person errors and gender errors in terms of the ease of association of form and function. Person-marking (especially first and second person) relates to stable, concrete, highly salient and personally relevant focal components of mental models. That should facilitate the formation of form-function associations. Gender in Russian is strongly related to phonological form (in the case of non-human

nouns) and to sex (in the case of human nouns). If it were only related to sex, we might expect the form-function association to form more readily than it does, as the sex of participants in mental models is also a concrete and relatively time-stable property. (But see Chapter 4 for discussion of my personal struggle with the acquisition of Russian pronoun gender.) Gender of inanimate nouns may play a role related to the lexicon and lexical access. Gender of an adjective could constrain noun expectations. If nouns are primed by their gender agreement, then verbal gender agreement in discourse might help to maintain a heightened level of activation of a noun that was used to introduce a participant. Discourse functions in general are presumably more complex than concrete referential functions, since dealing with a stream of changing mental scenes and interrelating later scenes to earlier ones is inherently more complex than dealing with individual scenes in isolation. Concrete referential functions come into play in the latter case, including ostensive reference in the here-and-now context, while discourse functions do not, at least not to the same extent (Thomson, 1996). This again should lead one to expect a difference in the ease of learning person agreement as opposed to gender agreement (in the case of inanimate noun subject gender agreement, at least).

Like verbal gender agreement, aspect-marking also faired badly in Experiment 2. The imperfective-for-perfective substitution faired worse in Experiment 2 than in Experiment 1, even though overall, participants in Experiment 2 detected a higher percentage of errors during the two listening tasks than did participants in Experiment 1 (the differences in means for the two experiments are significant for the Meaning-Oriented task, $t[89] = 2.252$, $p = .027$, two tail; for the Form-Oriented task $t[89] = 2.080$, $p = .040$, two tail; but not for the Printed Sentences Task $t[89] = .717$, $p = .476$, 2-tail). For the participants in these experiments, the processing of aspect inflection thus appears to be less active in general than the processing of case inflection. It is clearly less active than the processing of person agreement inflection and comparable to gender agreement processing. This may support the idea that aspect has a discourse function rather than a concrete referential function.

3.2.3.3. Further (qualitative) consideration of the specific errors

We will now consider the individual errors in more detail from a qualitative perspective. First, consider the **accusative-for-nominative** substitutions:

(11) Potom *chelovek-a stal myt' ruki pod kranom.
then person-acc started to.wash hands under tap:instr.
Then the man started to wash his hands under the tap.

Base-form: chelovek Correct target: chelovek Substituted: chelovek-a

(12) Cherez nekotoroe vrem'a *muzhchin-y prosnuls'a, vstal,
after some time man-acc awoke stood.up
After awhile, the man woke up, stood up,

podoshol k televizory i vykl'uchil ego.
went to TV and turned.off it
walked to the television and turned it off.

Base-form: muzhchin-a Correct target: muzhchin-a Substituted: muzhchin-y

As (11) and (12) involve the substitution of a non-base-form for a base-form (assuming nominative singular to be the base-form), in line with the results of Experiment 1 we might expect to see a reasonable level of sensitivity to this error. These two accusative-for-nominative substitutions involve the first and second declensions (respectively). Example (11) involves the addition of a syllable and (12) involves the alteration of a vowel. On the scale applied earlier, that means they differ in phonological salience: a rating of four for (11) and a rating of two for (12). Nevertheless, the sensitivity to the two errors is comparable, both in the Meaning-Oriented task and the Form-Oriented task. There were sixteen detections of the error in (11) and fifteen detections of the error in (12) in the Meaning-Oriented task. There were twenty-eight detections and thirty-one detections, respectively, in the Form-Oriented task. This suggests that declension, phonological differences and contextual differences did not have detectable effects on these two functionally parallel errors. Animacy and humanness were held constant. In fact, the two words are closely related semantically and in these contexts readily interchangeable. Their subjecthood is strongly semantically determined, providing a good opportunity for the detection of nonconverging cues. (*Chelovek* is somewhat more frequent than *muzhchina*, being ranked 64 and 643 respectively by Brown, 1996, in a ranking of 10,000 high frequency Russian words.)

Turning to the **nominative-for-accusative** substitutions we have (13) and (14).

(13) v spal'ne chelovek podoshol k krovati, vykl'uchil
in bedroom person went to bed turned.on

*lamp-a chtoby luchshe videt'.
lamp-nom in.order better to.see
In the bedroom, the man went to the bed, and turned on the lamp in order to see better.

Base-form: lamp-a Correct target: lamp-y Substituted: lamp-a

(14) Potom on podn'als'a, sn'al *rubashk-a i ostals'a.
then he rose took.off shirt-nom and remained

v br'ukax.
in trousers
Then he got up, took off his shirt, and remained in his trousers.

Base-form: rubashk-a Correct target: rubashk-y Substituted: rubashk-a

These errors are highly similar. Both involve second declension nouns. Both have inanimate referents and semantically plausible patients of the specific verbs, making them good candidates for direct objecthood, thus allowing a good possibility for the detection of a cue clash. Again, the levels of sensitivity to these two errors are comparable (twelve and thirteen errors, respectively, in the Meaning-Oriented task; thirty-four and thirty-one in the Form-Oriented task).

Turning to the **oblique-for-oblique** substitutions, we have (16) and (17).

(15) Posle togo kak namylil ruki *myl-e on polozhil ego
after that how he.washed hands soap:loc he put it

v myl'nitsu.
into soap.dish
After he soaped his hands with the soap, he put it into the soap dish.

Base-form: mylo Correct target: mylom Substituted: myle

(16) Kogda on lezhal v *krovat-ju, on popravil podushku pod
when he lay in bed-instr he adjusted the.pillow under

golovoj.
head
As he lay in the bed, he arranged the pillow under his head.

Base-form: krovat' Correct target: krovati Substituted: krovatju

Items (15) and (16) were included in an effort to correct possible problems with Experiment 1. Item (15) in Experiment 1 followed immediately after another error-containing item, and therefore, there was concern that the decreased detectability of the error in (15) was an artifact of the experiment. In Experiment 2, the text segment immediately prior to (15) did not contain an error, and there turned out to be no depression of the level of sensitivity to this error such as was apparently seen in Experiment 1. We observe here less detections of this oblique-for-oblique substitution than we see with the two accusative-for-nominative substitutions. In the Meaning-Oriented-Task, there were eleven detections for (15) vs. sixteen and fifteen detections for the two accusative-for-nominative substitutions. In the Form-Oriented-Task, there were twenty-three detections for (15) vs. twenty-eight and thirty-one for the two accusative-for-nominative substitutions. Taking accusative case, like oblique cases, to be marked, we might tentatively expect the accusative-for-nominative detectability to be partly parallel to the oblique-for-nominative detectability. That is, we might expect to find more sensitivity to accusative-for-nominative errors than to oblique-for-oblique errors. The observed difference in the number of directions consistently lies in the right direction, but the difference is not significant.

In contrast with (15), we were not able to change the general earlier picture of relatively high sensitivity to item (16) by modifying the experiment. Item (16) continues to stand out as a relatively highly detectable error, despite the fact that it involves an oblique-for-oblique substitution. It is the most frequently noticed case-error in the Meaning-Oriented task and in the Form-Oriented task, it is closely comparable to the other case errors, with the exception of the one in (15). Especially under the more demanding conditions of the Meaning-Oriented task (the task most like real-life listening, as well) the results of Experiments 1 and 2 suggest that this is a relatively robust error for the experimental participants. This fact does not readily yield to the types of functional explanations that seem compatible with the remaining items. It appears that a separate explanation is required here. Possibilities were explored in the discussion section of Experiment 1.

Turning to the verb inflectional errors, one was repeated from Experiment 1 and two were new. The one from experiment one, example (17) involving **incorrect aspect**,

was included to correct a problem which arose in Experiment 1 when a participant identified an identical error out loud in a group of seven participants.

(17) Potom on *zakr-yva-l glaza i usnul.
then he close-impf-past eyes and went.to.sleep
Then he closed his eyes and went to sleep.

Correct form: zakry-l Substituted form: zakr-yva-l

Although overall, both the Intermediate and Advanced participants in Experiment 2 detected more errors than the corresponding groups in Experiment 1, this was not reflected in the number of responses to (17). It turned out in Experiment 2 that there was relatively low sensitivity to this error (9.3% of participants in the Meaning-Oriented task and 22.2% in the Form-Oriented task; cf. 24.3% and 48.6%, respectively, in Experiment 1). At this point, the tendency of participants at the proficiency levels represented in this study to implement aspect cues during listening comprehension would appear to be relatively low, as noted earlier.

We can compare and contrast sensitivity to aspect cues with sensitivity to **gender cues** and **person cues** respectively in the cases of (18) and (19).

(18) On vyter ruki polotentsem i
He wipe hands towel:instr and

voda *peresta-l kapat' s ruk
water stop-past:masc drip from hands:gen:pl
Last of all, the man wiped his hands with the towel.He wiped his hands with the towel and the water stopped dripping from his hands.

Correct form: perestala Substituted form: perestal

(19) Chelovek *sidish v kresle i smotr'it telivizor.
person sit:2pers.sg in chair and watch:3pers:sg TV
The person is sitting in the chair watching television.

Correct form: sidit Substituted form: sidish

Item (18) contains the error to which the least sensitivity was shown in Experiment 2 and item (19) contains the error to which the greatest sensitivity was shown. In the comparison of case errors we saw that Experiment 2 was unable to detect statistically significant differences in sensitivity to different errors or error types. In this connection, the responses to verbal errors are methodologically encouraging, in that they demonstrate that in principle, the experiment did have the potential for detecting differences in sensitivity to different inflectional errors. Both examples (18) and (19) involve agreement. Gender agreement cues appear to acquire their functions much less readily than person agreement cues, as discussed above.

3.3. GENERAL DISCUSSION

The Meaning-Oriented task and Form-Oriented task are open to two important potential criticism (no doubt among others). We might question the extent to which participants in responding to stimuli with errors were actually reacting to the errors within

the stimuli, and not some other details. It was foreseen that this would be true to a certain extent. It was surprising, however, to see how many participants did not comply with the instructions to only mark definite, obvious errors about which there was no doubt in the Meaning-Oriented task., especially in Experiment 2. Four findings offer some reassurance that participants were to a reasonable extent detecting and reacting to errors in both listening tasks. First of all, there were relatively few cases where participants indicated errors in the listening tasks and were subsequently unable to identify the errors in the printed sentences task (in Experiment 1, less than 3% of detections in the Meaning-Oriented task, and less than 5% of those in the Form-Oriented task; in Experiment 2, approximately14% in both tasks). Second, there was a relatively low tendency not to re-mark an item in Form-Oriented task that was marked in the Form-Oriented task (11.4% in Experiment 1 and 11.9% in Experiment 2). Third, in both tasks of both experiments there were roughly twice as many reactions to items containing errors than to items without errors. The only obvious explanation for this is that participants were frequently detecting errors. Finally, the fact that non-random patterns of sensitivity to error-types emerged in other experiments, and that those patterns followed linguistically interesting lines, suggests that whatever error was introduced into the data by participants guessing, it was not enough to totally conceal the effects of the variables of interest (which includes the effects of group and task, in addition to the patterning of sensitivity to errors).

With respect to the effect of groups, it appears that the use of time since onset of Russian learning as a proxy variable for proficiency level appears to have been successful. This would not be a wise case in every situation. However, in a second language context where participants are heavily involved in regular use of the target language, it is not unreasonable to expect that a clear difference in mean learning time would show an effect of L2 development.

With respect to the effect of tasks, it seems quite possible that this reflected genuine differences in the types of processing involved. First of all, the performance of native controls reflects various interesting features of the stimuli. Regarding the ability of the Meaning-Oriented task to force meaningful processing and prevent metalinguistic (grammatical, analytical) reflections, even some native speakers commented on how demanding the found the pace of the Meaning-Oriented task. Next, there was almost no variation in the tendency of native controls to correctly identify all errors, and the few cases of their reacting to non-errors in Experiment 1 is possibly explainable in terms of a lexical problem (whether the man was *sitting* or *lying* at one point). This suggests that, at least from the native perspective, the errors were robust and not subtle in any way. Finally, there was almost no variation in the tendency of native controls to always select the correct pictures.

Turning to the experimental participants, the results of Experiments 1 and 2 are reasonably similar in the patterns of differences found in comparisons of groups and tasks. The overall higher performance of participants in Experiment 2 involved the Meaning-Oriented-Task and the Form-Oriented-Task, but not the Printed Sentences task. Experiment 2 was somewhat contaminated by the number of false alarms, which would tend to narrow the differences between errors that were more detectable and errors that were less detectable.[21] Yet large differences did emerge in the verbal inflectional errors. In

[21] The number of false alarms was a surprise, since it represented noncompliance with the instructions, particularly in the Meaning-Oriented task. In that task participants were asked to note "mistakes" only if they were "obvious", but to concentrate on the meaning, and on finding the correct pictures. Fortunately, it appears that the actual error detections were statistically strong enough to be detected in spite of the false alarms, both in the effect of groups (the Intermediates always had more false alarms than the Advanced participants, and yet the Advanced participants reacted to significantly more

any case, Experiment 2 did not yield the same clear patterns of differences of reactions to errors of case marking that were seen in Experiment 1.

In the remainder of this section I first discuss the implications of the more general findings: the low performance in relation to native controls, the differences in results between the two participant groups and the differences in results among the three experimental tasks. Following that, I seek to discern possible patterns in sensitivity to errors in the combined experiments, and to provide a preliminary account of them.

3.3.1. The pace of development

Nonnative users of Russian appear to process spoken input without consistently processing inflectional form. In the beginning, nonnative listeners may construct mental models by relying on lexical cues, word order cues, etc., with little implementation of inflectional cues for many months. There is a gradual increase in sensitivity to inflectional form over the course of a few years, consistent with what we would expect if inflection were being processed by learners increasingly over time. At least some form of inflection-related learning must occur, which in turn suggests that the processor has at least occasionally experienced inflectional form and comprehension processes simultaneously in working memory, allowing some aspects of comprehension processes to associate with some inflectional forms. Insofar as comprehension processes are insensitive to inflectional form, there is no basis for learning to occur. Therefore, the very fact of learning implies that at least occasionally the processor reacts to the presence of inflectional form.

As sensitivity to inflectional form increases, is the learner on a path leading in the direction of nativelike inflectional processing? To answer in the affirmative, we would like to see theoretically meaningful patterns in the developmental picture. The division into Intermediate and Advanced groups was done in a pseudo-longitudinal spirit. However, this aspect of the experiment revealed nothing regarding developmental sequences in the use of inflection during comprehension. Rather, similar relative tendencies were seen in both groups, only stronger in the Advanced Group than in the Intermediate Group. It appears that the direction of differences in sensitivity to various inflectional categories may be relatively constant over a long period of time, with increased sensitivity developing at different rates, but across the board, rather than mainly in one category and then mainly in another and so on. Rather than speaking of sensitivity to inflectional categories as developing *earlier* or *later* for one category than for another, we might speak of it as developing *less readily* or *more readily*. The readiness with which sensitivity develops for some inflectional category would presumably reflect the ease with which form-function associations are formed.

3.3.2. Sensitivity to case-marking

In the discussion of Experiment 1, it appeared that two factors needed to be posited in order to account for the varying degrees of detectability of the various error types. The first was the tendency of an inflection to trigger processing independently of other cues, and the second was the tendency of an inflectional cue to be expected in a particular context. It was suggested that oblique cases relatively frequently trigger processing regardless of other cues. This could be behind the relatively *moderate(intermediate) tendency for oblique-for-oblique substitutions to cause*

---

errors), and in the distribution of error detections by types. These facts also virtually rule out the attribution of the level of error detections to a response bias (Birdsong, 1989).

processing glitches. By contrast, it was suggested that nominative case does not trigger processing absolutely. Encountering a nominative case-marked noun in an oblique context had a relatively low tendency to cause the L2 processor to attempt to assign subjecthood (either in terms of semantic roles or discourse functions). However, when other cues led to the assignment of subjecthood, the noun to which subjecthood was assigned was expected to be in the nominative form (the base-form). Combining this contextual expectation with the tendency of oblique marking to trigger processing, we have an account of the *relatively strong tendency for oblique-for-nominative substitutions to be noticed.* At the other extreme, there was a *relatively weak tendency for Nominative-for-Oblique substitutions to cause reactions.* An explanation for this was proposed involving a hypothesis that in oblique contexts there is not a strong expectation of an oblique case, and nominative case does not trigger case-related processing (that is, nominative is not yet an independent cue to subjecthood or agenthood). The worst violations were the ones that involved both independent triggering by the encountered case form and a violation of case expectations (the oblique-for-nominative errors). The weakest violations were those that involved neither independent triggering nor a violation of expectations (nominative-for-oblique errors).

These two variables appeared at first not to account for the high sensitivity of participants to one of the oblique-for-oblique substitutions seen the Meaning-Oriented task in Experiment 1, a finding more or less replicated (though with less statistical clarity) in Experiment 2. It was pointed out that, within the confines of the errors in these experiments, neither the impossibility of the specific collocation involved in the error, nor the existence of prepositional government in the context provided an explanation by itself. It was suggested that a better possibility was simply that the specific correct target phrase is likely to be already familiar to participants. That is, common phrases could create a high contextual expectation of specific oblique forms. If so than the error *v krovatj-u*, 'in bed-instrumental' (as opposed to the correct *v krovat-i*, 'in bed-locative') could involve both the absolute triggering brought about by the oblique (instrumental) form, and an expectation of the locative form brought about by this specific phrase. That would make this error analogous to the oblique-for-nominative errors.

Experiment 2 added accusative case to the mix. Substitutions of nominative for accusative and vice versa resulted in similar numbers of detections, very unlike substitutions of oblique for subject and vice versa in Experiment 1. We might look for an explanation in which nominative case plays a stronger role in the nominative-accusative axis than in the nominative-oblique axis. Short of that, on the basis of the relatively high detectability of nominative-for-accusative errors, we would be led to say, in line with the approach embodied in Table 7, that the direct object context has a tendency to create an expectation of the correct case (since nominative case is relatively inert outside of its proper context). Does accusative case also trigger processing independently of other cues? If so, then accusative-for-nominative errors would be like oblique-for-nominative errors. In which case we would expect them to be the most detectable errors of all in Experiment 2. As they are not clearly more detectable than the other case errors, we will tentatively assume that for learners at the level represented by our experimental participants, accusative case, like nominative case, does not yet tend to trigger processing out of its proper context.

Since accusative case is the only new case in Experiment 2, we are now in a position to modify Table 7 as Table 11. This is a weak conclusion, as it is based on the failure to reject the null hypothesis regarding differing levels of sensitivity to errors of case marking. In addition, it strikes me as intuitively somewhat suspect. Nevertheless, in what follows, I will explore the consequences of accepting this conclusion, given its consistency with the experimental findings.

Error type:	Tends to violate case expectations that result from other cues?	Tends to triggers case-related processing?
Nominative-for-oblique (relatively low sensitivity on the part of participants)	**No** (oblique contexts appear not to create strong expectations)	**No** (nominative case apparently not tending to trigger processing independently)
Oblique-for-nominative (relatively high sensitivity)	**Yes** (nominative case tending to be expected when other cues indicate that a noun is subject)	**Yes** (oblique cases are somewhat prone to trigger processing independently)
Oblique-for-oblique (intermediate sensitivity)	**No** (established above)	**Yes** (established above)
Accusative-for-Nominative (intermediate sensitivity)	**Yes** (established above)	**No** (accusative case not appearing to tend to trigger processing independently)
Nominative-for-Accusative (intermediate sensitivity)	**Yes** (accusative case tending to be expected when other cues indicate that a noun is direct object)	**No** (established above)
Oblique-for-Accusative (prediction: high sensitivity)	**Yes** (established above)	**Yes** (established above)
Accusative-for-Oblique (prediction: low sensitivity)	**No** (established above)	**No** (established above)

Table 11. Two hypothetical factors extended to account for differing levels of sensitivity to different types of case errors in Experiments 1 and 2.

In Table 11 there are three ways that a combined "yes" and "no" in a single row might be interpreted with respect to the ranking of that row in relation to other rows. We might rank a yes-no column in the table more highly than a no-yes column. We might rank a no-yes column in the table more highly than a yes-no column. Or we might rank a yes-no column and a no-yes column the same. Suppose we say that all rows with "yes" in the first column are ranked higher than any row with "no" in the first column. That amounts to saying that strong expectations are more important than independent triggering. With that added assumption, Table 11 predicts the order of sensitivity to case errors shown in Table 12.

I. Yes-Yes: oblique-for-nominative
II. Yes-No: accusative-for-nominative, nominative-for-accusative
III. No-Yes: oblique-for-oblique
IV. No-No: nominative-for-oblique

Table 12.. Order of sensitivity to case errors predicted from Table 11, from most to least. This includes the assumption that Yes-No is ranked higher than No-Yes. The "Yes" and "No" values are those in Table 11.

Note the assignment of "yes" or "no" to a given case context or case inflection was motivated. The hypothesis that nominative case is inert in non-nominative contexts, the hypothesis that oblique contexts do not strongly expect oblique cases and the hypothesis that oblique cases have some tendency to trigger processing independently were motivated by rejecting the null hypothesis regarding the distribution of error detections in Experiment 1. The hypothesis that direct object context creates an expectation of accusative case and the hypothesis that accusative case is not a strong independent processing trigger were motivated by our inability to reject the null hypothesis (with respect to case inflection) in Experiment 2. As noted, accepting the null hypothesis provides at best a weak motivation. However, the strong differences in detections of verbal errors in Experiment 2 suggested that the experiment did in fact have the power to detect clear differences that were present. As to the ordering of sensitivity to errors which Table 11 predicts, II and III in Table 11 have not been shown to differ empirically. They can be reduced to a single level of sensitivity by abandoning the hypothesis that yes-no overrides no-yes in the ranking. However, there was a non-significant difference in the direction of the ranking shown in Table 12, provided we also accept the hypothesis that highly familiar phrases can create strong case expectations. That means that one of our oblique-for-oblique errors, the one in example (17) (= example 7) fits the pattern of Table 11 and the other, in example (16), lies in the right direction, but not significantly so. The data thus appear to be internally consistent in terms of the two factors hypothesized (activeness of the specific category of case inflection and strength of contextual expectation of a specific case) and the predicted effects of those factors.

The hypotheses embodied in Table 11 make predictions regarding errors which have not yet been tested: those which would result from substituting accusative for oblique cases or vice-versa. The predictions are that an oblique-for-accusative error would be relatively highly detectable (yes-yes), on a level with oblique-for-nominative errors, while an accusative-for-oblique error would be relatively low in detectability (no-no) on a level with nominative-for-oblique errors. The first prediction has more intuitive appeal to me than the second.

The property of creating the expectation of a specific case-marking (ignoring the issue of highly familiar phrases) coincides here with the property of being one of the focal participants (subject or object). The property of being an absolute trigger coincides here with the property of being a case marking for non-focal participants. In other words, the specific semantic roles assigned by oblique cases and direct cases may play a smaller role than the discourse roles assigned.[22] Semantic role assignments are only clearly implicated by the sensitivity to oblique-for-oblique errors.

---

[22] As noted earlier, the discourse role of case is already involved in single sentence utterances, unlike the discourse function of verbal aspect. Specifically, it involves assigning (or maintaining the assignment of) the primary focus of attention in a mental model to a particular token.

In order for nativelike proficiency to be achieved, there would have to come a day when all contexts create appropriate expectations and all case forms trigger processing independently (allowing for various complications created by inflectional homonymy and the like).

The picture of L2 Russian case-acquisition that is taking shape here is as follows. At the outset the learner does not process case at all. Both discourse roles and semantic roles are assigned on the basis of lexical semantics, word order, the plausibility of the mental model constructed, and perhaps other cues, but not case cues. Two important steps in mental model construction involve the merger of the primary and secondary focal participants into the model. The primary focal participant is generally highly accessible in memory and thus would not benefit so much from special marking, that is, from deviation from the base-form. Most often this participant will be high on the animacy hierarchy and will occur early in the sentence. Such a noun will be chosen as subject unless the processor is told to do otherwise, that is, unless there is some clear phonetic deviation from the base-form of the word. Therefore the base-form itself would rarely need to trigger special processing beyond the work done by other cues. As long as the subject is determined, determining the direct object is also possible without reference to case cues in most instances. The subject-assignment step in comprehension and, to a lesser extent, the object-assignment step, are ubiquitous steps in the process of mental model construction, creating a relatively good opportunity for form-function associations to take hold. However, what appears to develop initially is a context-based expectation, and not the independent activeness of the case cue itself. Once the subject and object are determined during the processing of an individual sentence, the task of determining oblique roles is greatly simplified. In our example (2) above, repeated here as (20), there are two potential distinct direct objects. However, as hands make an excellent choice as to what people wipe after having washed their hands, its object status can be readily determined. In that case, given the role of towels in washing hands, the model is going to treat the towel as an instrument, barring good reason to do otherwise. In the case of most other oblique noun types, the semantic role is partly marked by a preposition (an example of distributed exponence—see Chapter 2).

(20)	Nakonets	chelovek	vyter	ruki	*polotentse.
	Finally	person:nom	wipe	Hands:nom/acc	towel:nom/acc

Last of all, the man wiped his hands with the towel.

An oblique noun of any given type is a comparatively rare occurrence compared to subject and direct object nouns. Thus the expectation of oblique case marking, given the independently determinable nature of the semantic role, will be slow to develop. On the other hand, any oblique case marking functions as a processing trigger (initially, perhaps, contributing mainly to the assignment of non-focal participant status) on the basis of its non-identity with the nominative and accusative forms.

In summary, such considerations suggest that the pattern of distributions of error detections observed can be understood in processing terms. Sensitivity to errors is a function of expectedness of case-marking and the ability of case-marking to trigger processing, both of which factors should have their place in any account of language processing.

3.3.3. Sensitivity to verbal inflection

In terms of verbal inflections, it appears that aspectual cues are slow to get off the ground in L2 Russian development. This is perhaps more compatible with the discourse

account of aspectual function than an account in which aspect is taken to express temporal semantics. If the function being acquired were one of marking the temporally extended nature of an activity or the punctuality of an action, these would be moderately concrete referential functions. By contrast, backgrounding and foregrounding are purely discourse sequencing and staging functions. I have argued elsewhere (Thomson, 1996) that concrete referential functions are inherently easier to acquire than discourse sequencing or discourse staging functions. Processing continuous discourse requires the ability to deal with individual conceptual scenes plus the ability to sequence them. It would seem that dealing with individual scenes would have to be easier than both processing individual scenes and sequencing them. Referential semantics is involved in dealing with isolated individual scenes, while discourse semantics is not (with the caveat that the discourse functions of case marking play a role already in single sentences utterances).

A major role of verbal gender agreement might be topic maintenance in discourse. That is, once a (let's say) feminine primary topic of discourse is established, ongoing feminine subject agreement (together with the absence of full noun phrase subjects) would be a signal that no change in primary topic has occurred. The listener thus has no reason not to keep the "spot of light" (focal center) on the same primary character. This is again a much less concrete semantic function than person agreement which commonly functions to identify which speech act participant is to be merged as agent into the concept partly expressed by the verb. Gender agreement in Russian, if not generally, thus relates more commonly to displaced narrative discourse, while person agreement relates is heavily involved in here-and-now discourse. This may help to account for the apparent fact that the mental associations involved in the acquisition of person cues appear to form more readily than the mental associations involved in the acquisition of verbal gender agreement cues.

We will have more to say about the possible implications of Experiments 1 and 2 in the following chapter. Hopefully the preceding discussion suggests the potential fruitfulness of approaching the grammatical form of a language in general, and an inflectional system in particular, as a set of processing cues that must be acquired through the association of patterns of speech with events involved in the comprehension process.

4. The Long Road Toward Full Inflectional Processing

In the introduction I posed three questions:

(1) Do adult learners of L2 Russian develop the ability to make use of inflectional morphology in on-line comprehension?

(2) Is there a pattern in the development of sensitivity to inflection, and if so, what is its nature?

(3) What, if any, is the relationship between metalinguistic (analytical, formal grammatical) strategies and the sensitivity to inflection that is evidenced in on-line listening?

These questions were of interest, among other reasons, because Russian inflection is be an area in which nativelike production by L2 users is not commonly. That is, it is an apparent area of so-called fossilizable linguistic features. Part of the original puzzle of fossilization was the tendency of learners to revert to nonnative patterns after the achievement of some degree of nativelikeness in spoken production. I suggested that if we were to look at comprehension processes that make use of fossilizable features, rather than only looking at spoken output, we might find a more monotonic pattern of development. More specifically, some of the progress and backsliding seen in spoken output might turn out to be due to the employment of strategies such as metalinguistic planning, formulaic language, the reuse of specific phrases and so on. That is, in terms of providing a guide to the linguistic aspects of L2 development, production data are probably somewhat clouded by these other varieties of strategies. Of course, the contributions of various factors to the form of spoken production also need to be sorted out, but the potentially simpler developmental pattern that might be discovered in the comprehension system could contribute toward that end, since the reactions of the comprehension system must be one of the important factors influencing the production system.

The broad framework chosen for examining developments within the comprehension system (Chapter 1) was primarily taken from the psycholinguistic literature on language comprehension. Rather than looking at L2 knowledge as static knowledge of grammar, that is, knowledge of allowable arrangements of linguistic elements, we chose to look at linguistic elements as processing triggers, and thus to view acquisition as a matter of coming to react to processing triggers in particular ways.

Against the background of such assumptions, we found evidence in Chapter 3 related to the three questions raised in the Introduction. That is, we saw that (1) inflection-related learning appears to be going on in our L2 learners (sensitivity to cue clashes was higher for Advanced participants than for Intermediate participants); (2) there is some evidence of patterning, that is, of systematicity, in the developments (there was a non-random distribution of reactions to different cue-clash types, along theoretically reasonable lines); and (3) the relationship of metalinguistic knowledge of linguistic elements to the exploitation of those elements in L2 comprehension may be complex, but it appears that the development of comprehension processes have as much of an impact on the acquisition of metalinguistic knowledge as vice-versa. First of all, when Intermediate participants displayed a high level of ability to apply metalinguistic analytical knowledge in the Printed Sentences task, they did not display a correspondingly high

level of sensitivity to cue clashes in the listening tasks. Second, the Advanced participants, with higher proficiency over all, showed fairly consistent ability to apply metalinguistic analytical knowledge of the sort that most of the Intermediate learners had presumably also been taught (and drilled on) from an early point in their learning. Yet, the Intermediate learners evidenced some large gaps in their ability to apply such knowledge even to the analysis of simple written sentences.

Returning to question (1) above, although there appears to be progress over time towards more nativelike use of inflectional morphology, there is little evidence that might assure us that learning would ever lead to nativelike or near-nativelike processing. In this final chapter, I will add a case study that at least opens the possibility that with massive enough experience in the language, performance (in comprehension) might continue to become more and more nativelike.

In relation to question (2) above, I further argue in this chapter that the pattern of learning suggested in the previous chapter fits well with a somewhat appealing learning scenario that places inflection and inflectional learning within the lexicon (in the general spirit of Bybee, 1985). I also take some liberties in this chapter in sharing introspective observations of my own experiences in comprehending L2 Russian, noting apparent difficulties in inflectional processing. These reflections may cast further light on the dynamics of comprehension that were tapped by the experiments reported in the previous chapter.

Two broader conclusions are suggested by the findings in Chapter 3. The first involves the time-frame of L2 Russian inflectional acquisition or, closely related to that, the volume of experience with the target language that appears to be necessary for the moderate to thorough acquisition of many aspects of the inflectional system. The necessary volume of experience with the language appears to be truly enormous. And yet many learners are likely to feel that with regard to the inflections examined in Chapter 3, and the uses of those inflections that were examined, their spoken production is more nativelike than would appear to be reflected in their use of those inflections in comprehension. This seeming paradox is discussed in relation to proposals regarding the role of metalinguistic knowledge in L2 spoken production. The findings of Chapter 3 probably support those views which distinguish between normal linguistic functioning and the use of what Schwartz (1993) called learned linguistic knowledge.

The second broader conclusion is also related to the apparently huge volume of experience required for the thorough acquisition of many aspects of L2 Russian inflection: the process of learning to exploit inflection during comprehension may be an extremely gradual one. The picture that started to emerge in Chapter 3 (to be further supported by the case study and introspective reflections of this present chapter) could cast some doubt on any idea that learning involves categorial changes to linguistic competence system resulting in knowledge that is invariant at each point along the way, with observed variation being attributed to performance mechanisms or to temporary competing hypotheses within the competence system. Rather, variation with respect to specific elements could be inherent to the learning mechanism, since learning may involve gradual strengthening (both in comprehension and production) and not occasional leaps, and factors influencing production at any moment may be multiple.

In connection with these two broader conclusions, as well as with the more basic findings of Chapter 3, a few current models of second language learning are given brief consideration in this chapter.

I suggest that existing models tend not to place adequate responsibility for learning within performance systems. Evidence of learning, in most models, is sought in the patterns that emerge in the output of the production system. The development of processing mechanisms, with comprehension mechanisms taking some logical priority

over production mechanisms, is often taken for granted and not viewed as entailing learning specific to those mechanisms.

4.1. CASE STUDY: A MORE HIGHLY DEVELOPED L2 RUSSIAN USER

As noted, it might appear at this point that we are justified in wondering whether L2 Russian users in the population represented by our experimental participants will ever reach the point of consistent nativelike functioning.

It would be interesting to study a sample from the population of nonnative Russian users who began learning Russian as adults and for whom Russian has become the dominant language of everyday life. Unfortunately, such individuals did not prove easy to find, perhaps because the Soviet Union was not a major immigrant-receiving nation. Most of the participants in my experiments saw themselves as living in Russia temporarily and tended to socialize heavily with others of their own native language backgrounds, or even to use L2 English as a lingua franca in preference over Russian whenever possible. While all of the experimental participants were also using their L2 Russian in major ways in their everyday lives, it is doubtful that any would have considered Russian their dominant language.

In the course of my search for participants, I located only one nonnative speaker who had settled in Russia permanently and adopted Russian as her dominant language. I contacted her after Experiments 1 and 2 had been completed, and asked her if she would be willing to perform the tasks from Experiment 2 and to discuss the experience with me.

In the interview, this woman (who happily waived the anonymity clause in the experiment consent form) credibly claimed that Russian was her dominant language and that it had enjoyed that status for perhaps four years. She grew up in Ethiopia and Finland. In early childhood she spoke mainly Finnish and English and some Amharic. At the age of fourteen, she received limited formal instruction in Russian as a school subject in Finland for one academic year, amounting to perhaps three hours of exposure per week. At the age of twenty, ten years before I interviewed her, she moved to Russia in hopes of attending medical school. Like many of the experimental participants, she began her serious Russian learning in a preparatory program that was designed to enable learners to participate in university courses following an intensive eight months of Russian training. After completing the preparatory Russian program, she went on to complete one and a half years of medical school in Russia. (At that point her profile would have been quite similar to many experimental participants in the Intermediate groups.) Her career goals then changed when she married a Russian man and discontinued medical school. At the time of the interview, she had been married for seven years, all of which time Russian had been the primary language of her marriage relationship. She has a Russian-Finnish-English trilingual daughter whom she most commonly addresses in Finnish; otherwise her day-to-day life involves only minimal relationships with Finnish or English speakers. She and her husband pastor a Russian protestant church and she is frequently involved in pastoral counseling, in addition to carrying on a generally active social life with Russians on a day-to-day basis. Although she felt that Russian was her dominant language, she stated that she did not notice any difference in difficulty functioning in Russian, Finnish, or English. (I interviewed her in Russian.)

In performing the Meaning-Oriented task from Experiment 2, this woman reacted to all items containing errors. Unlike native speakers, she also marked one item as containing an error when it did not. Therefore, before she performed the Form-Oriented listening task, I made a further point of asking her to be extremely strict with herself and only indicate the presence of an error if she considered it to be beyond doubt. In

performing that task, she behaved qualitatively like native speakers I observed when hearing errors. That is, she immediately and without hesitation, marked an "X" on the form opposite the number of the appropriate item, not devoting noticeable time to reflection. She differed from native speakers in the case of items that did not contain errors. With those items, she frequently verbalized doubt, pondering aloud whether there might be an error, but not marking an "X", since she was not certain. In other words, she appeared to react automatically, on a language processing level, to true errors, but had a low level of confidence in her metalinguistic judgements regarding sentences without errors. This is perhaps not surprising. Immediate reactions to strong, clear cue clashes would be reflected in the case of items containing errors. However, in the other cases, an adult-onset L2 learner might often wonder whether there is "something I don't know about". There has been debate regarding the claim that L2 learners are less certain about the grammaticality of grammatical sentences than they are about the ungrammaticality of ungrammatical sentences (Felix, 1985; Birdsong, 1989). This woman's behaviour was consistent with that claim.

This woman's performance opens the possibility that it may in principle be possible for a nonnative Russian learner to develop nativelike sensitivity to inflectional cue clashes, including those involving verbal aspect. However, we must still bear in mind that the experimental errors were embedded in relatively simple texts, and they were errors that would be considered very elementary in nature in a pedagogical context. We do not know what we would have observed had we looked at a wider variety of contexts and functions of inflections. In addition, we do not know what portion of the population that this woman represents (people who learned Russian as adults and adopted it as their primary language of everyday life for several years) would show the same high level of sensitivity to inflectional cue clashes. Nevertheless, the picture she presents is consistent with the direction in which the experimental participants appeared to be moving. That is, the pattern seen in those participants was consistent with a plausible pattern of growth toward nativelike functioning. This woman, with presumably greater experience of Russian than most if not all of the experimental participants, showed further growth in the same direction.

4.2. PRIMITIVE DEVELOPMENT OF INFLECTION IN L2 RUSSIAN

In the previous chapter, there appeared to emerge a general picture of systematicity in the growth of sensitivity to inflection. The case study of the previous section appears to illustrate a stage further along the same developmental path. It is possible, of course, that the apparent pattern of sensitivity to inflectional form that was observed has nothing to do with the development of a truly linguistic inflectional processing system. To further support the view that such a system *is* (at least in part) gradually developing in (many) adult learners of Russian as a second language, I attempt in this section to refine and elaborate a plausible developmental scenario consistent with the findings of the previous chapter, focussing especially on case inflection.

Turning to that scenario, we might propose the following pattern of developments in line with the findings reported in Chapter 3. First, nouns are registered in the mental lexicon primarily in their nominative forms. Other forms are heard, but the processor treats them as exemplars of the (nominative) prototype and, in line with this, they do not trigger any extra processing related to the specific formal deviation of the exemplar from the basic lexical prototype (beyond whatever processing is needed to equate the encountered form with the prototype). At this stage in the development of many lexical items, variants of a lexeme would simply be equated with the learner's primary version of the word. A spin-off of this within the production system might at times be observable in

nonnative recasts of native utterances, such as the one in (1b), in which a learner was observed to substitute the nominative form of a noun for a native speaker's oblique form:

(1a) Russian waiter: S gaz-om ili bez gaz-**a**?
with gas-instr or without gas-**gen**
With or without carbonation? (Asked in reference to mineral water)

(b) Nonnative customer: *Bez gaz.
without gas:**nom**
Without carbonation.

Such faulty repetitions of native speech by nonnative speakers, replacing non-base-forms with base-forms, were commonly observed in nonnative Russian speech. (They might be considered naturalistic cases of elicited imitation.) An example involving an adjective is shown in (2), assuming that the base-form for adjectives is the masculine nominative singular form.

(2a) Native Speaker: Dlinn-**oe**? Krugl-**oe**?
long-neut:nom:sg round-neut:nom:sg
Is it long? Round?

(b) Non-Native speaker: Krugl-**yj**.
Round-masc:nom:sg
Round

Based on similar experiences of my own, I can imagine that the learner here had a sense of having heard *kruglyj* when he in fact heard *krugloe*. Hearing a non-base exemplar of the base-prototype might primarily strengthen the prototype form at the stage of development postulated.

The move to the next stage might be gradual, affecting different words at different rates. At this stage, the lexeme is well-established and thus easily processed. If enough instances of a particular non-base-form of a particular lexeme are heard, that non-base-form would develop some independence, and if independent enough (that is, if strongly enough represented, while not functionally differentiated), could start to compete with the base-form during speech production. In terms of a spreading activation model (Dell, 1986; Dell, et al., 1997; Dell & Seaghdha, 1992; Stemberger, 1998), the various inflectional variants of a lexeme could be represented in a lexicon, with different resting activation levels. Various factors would change the activation levels, priming one form or another. For example, having just heard a particular non-base-form could raise its activation level so that it would win out over the base-form during lexical retrieval, as in the interchange observed in (3).

(3a) Native Speaker: Kogo **ty** l'ubi-**sh**?
Who **you** like-**2pers**:sg
Who do you love?

(b) L2 speaker: ***Ja** nikogo ne l'ubi-**sh**.
I no-one not like-**2pers**:sg
I don't love anyone.

In example (3), the second person singular verb form (arguably not the base-form) is used in a first person environment as an apparent perseveration of the form used by the native speaker in the exchange. Such apparent perseverations were occasionally observed and could suggest that the offending form is well represented in the lexicon at this point of development. Nevertheless, it is still capable being equated with the target form in terms of function. Of course, to demonstrate that substitutions such as those seen in examples (1) and (2) are an earlier phenomenon than perseverations such as (3) would require substantial research.

Recapitulating, the first stage of development of sensitivity to inflection might be one in which the base-form is the prototype and other forms are treated as functionally equivalent exemplars. The second stage might be one in which the non-base-forms have relatively strong representations of their own, but are still not functionally distinct from the base-forms. (Keep in mind, this pattern of development would not require discrete stages. Rather, there could be gradual changes in resting activation levels as word-forms are repeatedly encountered in speech.)

Non-base-forms naturally carry the potential for triggering extra processing. The base-form is the default form and hence the most expected form. When non-base-forms occur in comprehended speech and succeed in activating non-base-forms in the learner's mental lexicon, there could be an added "surprise" value of the non-base-form, over and above its basic ability to trigger recognition of the lexeme to which it belongs. In the domain of case-marking, the "surprise" would always coincide with a noun that did not refer to the primary focal participant at the point in the discourse (that is, a noun other than the subject noun). This would allow the most primitive level of form-function association that may have been evidenced by certain aspects of the results in Experiment 1. In that experiment, we saw evidence that participants frequently expected the nominative case in subject position, but were far less likely to react to its inappropriate use in non-subject positions. The specific details of case form often appeared to be ignored when the noun was nominative-marked. Non-base-forms were treated as non-subjects. However, it was apparently expected that there will be no word-form "surprises" connected with the primary focal participant (grammatical subject).

A next phase might involve various oblique categories taking on more specific functions, such as marking instruments and locations. This further development is reflected in the fact that in Experiment 1, oblique-for-oblique substitutions had some tendency to trigger reactions, though smaller than the tendency for oblique-for-nominative substitutions to trigger reactions. Oblique case inflections would now have two functions: signaling non-focal status (non-subjecthood) and marking specific semantic roles. In the oblique-for-oblique errors only the second function would be violated.

The second experiment failed to detect a hypothesized earlier stage at which accusative case-marking triggers extra processing beyond that triggered by base-forms. If there is such a stage, then it will be evidenced by accusative-for-nominative substitutions triggering a stronger reaction than nominative-for-accusative reactions. However, in Russian the nominative and accusative case forms of a lexeme are often non-distinct. Thus, the base-form inflection is not a consistent cue to subject status. On the other hand, when accusative case *is* distinctively marked, it is quite a strong cue to object status (Kempe & MacWhinney, 1999). To further complicate the picture, the idea that direct objects refer to secondary topics (see Chapter 2) places them closer to the focus of attention than oblique nominals. In Experiment 3 such factors may have conspired in such a way that substituting nominative forms in object position and substituting accusative forms in subject position had similar effects. Kempe and MacWhinney (1999) interpreted their results as suggesting that the nominative-accusative distinction takes hold

relatively readily in L2 Russian, when comparison is made with L2 German. Our results do not appear to contradict that finding.

A further stage, not evidenced in the participants in the experiments reported in Chapter 3 (except perhaps in the case of the phrase *v krovati*, 'in bed') would be one at which not only do oblique cases trigger specific processing, but oblique environments also create a strong expectation of the appropriate cases. Although we did not see widespread signs of this development in connection with oblique forms in Experiment 1, we did interpret the evidence in Experiment 2 as indicating that direct object context had come to create an expectation of accusative case.

Summarizing, from a processing perspective, the following pattern of phases in the development of sensitivity to inflectional form (focusing here especially on case form) is reasonably plausible, and is not incompatible with the findings of the experiments reported in Chapter 3:

A. Base-form as prototype, other forms equated with it regardless of context.
B. Non-base-forms strong enough to compete with base-forms for selection in production, given adequate priming (but still not functionally distinguished).
C. Functional differentiation begins, as non-base-forms become associated with non-subject status; base-form expected in subject context.
D. Accusative form comes to be expected in direct object context.
E. Oblique forms become functionally differentiated from one another.
F. Specific oblique forms come to be strongly expected in specific oblique contexts and nominative case is now a strong, independent cue to subjecthood.

As noted, these would not be discrete stages, but rather a direction of drift in the noun lexicon as a whole, as base-forms, non-base-forms and links to functions are gradually strengthened. A certain level of strength of an early base-form might be required before non-base-forms would readily contrast with it. That would in turn allow the strengthening of the non-base-form to get underway. Once non-base-forms are adequately strengthened they can become functionally differentiated beyond their mere association with non-subjecthood.

When the comprehension system of an individual becomes so well developed that it consistently and strongly expects the contextually appropriate forms, we might expect that it will be hard for the production system to go on producing utterances which clash with those expectations, barring some special mechanism which prevents the nonnative speaker from comprehending his or her own speech. There could even be a direct relationship between the predictive proclivities of the comprehension system and the mechanisms of production. In any case, there must be a stage in lexical development when common inflectional patterns are generalized to lexemes for which specific required inflectional forms have not previously been encountered in comprehension (as demonstrated, say, by Berko's "Wug test"—Berko, 1958). In Bybee's (1985, 1991) lexical model, this would happen once many forms with the same ending have become associated with the same function.

In the case of Russian we saw that the picture is particularly complicated. Generalization of a pattern to new instances cannot be a matter of, let's say, *nouns ending in –u* becoming associated with *patienthood*. Rather the generalization must take the base-form into account: *nouns ending in –u and having base-forms ending in –a* become associated with *patienthood*. The development of associations among members of larger sets of inflectional forms connected by shared functions (such as the function of signaling patienthood, regardless of noun declension) would also be necessary on the

path to nativelike functioning. One bit of relevant evidence in Experiment 2 might conceivably be seen in the fact that an erroneous first declension noun and an erroneous second declension noun appeared to be quite similar in detectability. A consistent finding along these lines could hint that the category of accusative case is developing uniformly across different declensions (at the stage of development represented in the sample). Once developed, large networks of word-forms united by a common function would then be available for reuse in new functions (see Chapter 2) [23]

Turning to verbal inflection, in terms of the levels of sensitivity that we saw, sensitivity to person agreement appears to develop more readily than sensitivity to aspect marking or verbal gender agreement. It was not expected that there would be so little sensitivity to aspect errors, particularly in the case of the perfective-for-imperfective substitutions. However, the early advantage of person marking was expected. The person marking system is involved in deictic reference in here-and-now discourse situations and thus often relates to prominent and, more importantly, stable building blocks of mental models. The other verbal inflections, once they begin to be used contrastively, would arguably serve discourse functions such as indicating that an activity is background to an event, and that it does not increment the mental model to a new stage. Verbal gender agreement participates in the topic continuity system. Insofar as it is linked to the real-world sex of a referent, the opportunity for form-function associations to develop might be increased (but see the discussion below of my own struggle with pronoun gender).

All in all, we have seen nothing to suggest that sensitivity to inflection in L2 Russian is not evolving in learners along a plausible developmental path. We could hypothesize, on the basis of our limited evidence, that many of these L2 users are in the process of acquiring the inflectional processing system as part of their L2 Russian comprehension systems. However, the pace of development would appear to be such that full-acquisition must take a very long time. One of the most advanced participants, with several years of life in a heavily Russian milieu and a strong personal commitment to learning and using Russian, was insensitive, for example, to the substitution of *br'ukax*, 'pants:locative:plural' for the correct target *br'uki*, 'pants:nominative:plural'. He discussed that particular item after the experiment, and it appeared that there had been no attention lapse. He had heard the inflectional form, but it had remained inert, the locative form being treated as equivalent to the nominative form, even with this relatively fluent L2 Russian user. For this learner, the phases of development A through F discussed above would appear to be simultaneously in evidence across the inflected forms employed in the experiments, in line with the idea of gradual, across-the-system strengthening in the long-range direction of consistent nativelike processing, rather than clear (or semi-clear) successive steps in acquisition.

4.3. PERSONAL INTROSPECTIVE REFLECTIONS

In carrying out this research, I was in the not terribly common position as an SLA researcher of being a member of the population represented by the experimental participants. While conducting the experiments, I met the criterion for the Intermediate group. As an ongoing learner of Russian, I now belong to the population represented by the Advanced group: I have recently passed the four year mark from the onset of my Russian learning and I lived in Russian speaking countries for nearly two and one half years. Although the rather unconventional variety of evidence offered in this section must

[23] As noted, it may sometimes be the case that many adult learners are under pressure to produce target-like forms in a variety of functions long before such developments could have taken place within in their comprehension systems.

be treated with appropriate caution, I would feel I were being remiss were I not to share examples of my own introspective experiences related to the phenomena under investigation. I suspect I have developed a rather high level of introspective sensitivity to such matters, and my observations, to the extent that they are generalizable, could clarify some aspects of the performance of experimental participants in a way that is not often possible in a research context.

Recall that one of the least noticed errors in Experiment 1 was the one in which the imperfective verb *stojat'*, 'be standing' was replaced by the perfective verb *postajat'*, 'stand for a limited time period'. This was the only error involving a prefix. As noted, the correct target verb is an extremely high frequency lexical item and it appeared that the prefix was ignored by most participants, though it is hard to imagine it in general not being perceived. When listening to Russian texts myself, I do find evidence of prefixes being in essence ignored by my language processor. To cite a recent example, I heard the word *vyrabotat'*, 'work out' in a video about baby care. In context, what was being said was that a newborn baby quickly *works out* a regime of sleeping and waking. Moments after hearing *vyrabotat'*, I attempted to recall the sentence that contained it, specifically trying to recall how the idea translated 'work out' was expressed. I was confident that I had heard a prefixed form of the stem *rabotat'*, 'work' but could not recall the particular prefix. I guessed that the word had been *razrabotat'*, also translatable as 'work out'. I then rewound the videotape and found that the word was *vyrabotat'*. Upon hearing the word again, I had a clear sense of recognizing it as what I had heard the first time. It may be that I initially understood the word, relying on the root and the context, without making use of the prefix. In any case, it appears that the stem left a stronger memory trace than did the specific prefix. Perhaps the stem was used for lexical access (being itself a high frequency lexeme) and the specific prefix was not, although the fact that there was a prefix was also at least registered.

Although this example involves a prefix, analogous experiences with suffixes are common as well. Often I hear a noun which I (apparently) take to be nominative, but later in the sentence realize that it must have had another case value. Quite surprisingly, at times such a "slip of the ear" has occurred even with a pronoun (where the inflectional device is not suffixation, but rather suppletion). One such example occurred recently when I heard sentence (4).

(4) Emu kapel'nitsu uzhe sdelali.
him:dative IV.drip already did:3pers:pl
They have already started him on an intravenous drip.

When I reached the end of the sentence, metalinguistic reasoning led me to believe that I had made two "slips of the ear" on first listening. Grammatically, recipients of intravenous infusions, like recipients of invasive medical procedures generally (injections, vaccinations, surgical operations, etc.), are marked with the dative case in Russian (while the noun referring to the invasive procedure is the accusative-marked direct object of the verb *delat'*, 'do'). However, my L2 Russian language processor appeared to treat *emu*, 'him:dative' (the third person singular dative pronoun) as *on*, 'he' (the nominative counterpart). I had the sense that I was looking for a verb for which the referent of the sentence initial pronoun, i.e., participant at the center of the discourse (the narrator's son), would be the grammatical subject. Hence a glitch occurred when the verb *sdelali*, 'they did' was encountered. It had its own (unspecified third person plural) subject. I was highly aware of performing a reanalysis at that point. Fortunately, I was able to draw on my own short term memory for the reanalysis, since the form *emu* was still retrievable in

an auditory representation (even in the speaking voice of the narrator). That is, although I had understood *emu*, 'him:dative' as *on*, 'he:nominative', it seemed that I had perceived it as *emu* (the dative form) with full phonetic detail. This could well illustrate a partial dissociation of the specific phonetic form from the set of processes triggered by it. It could also illustrate the equating of an exemplar with the base-form prototype.

To continue, having reanalyzed *emu* as an experiencer, I was left with *kapel'nitsu*, 'IV.drip:accusative', as the direct object. I reasoned that it must have been in the accusative form, although I felt I had heard it in the base-form. I rewound the tape and heard a very clear *kapel'nitsu*, that is, the required accusative form. A further interesting detail occurred at this point. As I relistened to the sentence trying to hear both the form of the pronoun (*emu*) and the form of the following noun (*kapel'nitsu*), I found I could not easily hear the inflectional details of both words on the same listening. I could readily hear one or the other, but only after several listening could I hear both. (This might be taken as reinforcing a finding of VanPatten, 1990, that monitoring for grammatical morphemes interferes with processing). Note that on first hearing, I did in fact understand the sentence correctly and rapidly, although successful understanding appeared to involve a garden path and reanalysis. Had I been a less reflective L2 Russian learner (as, thankfully, I often am), I suspect that my comprehension of the text would simply have marched successfully onward, in spite of the brief problems occasioned by the failure of my language processor to fully implement the inflectional forms.

In the case of *emu* in example (4), on the account assumed here, two processing instructions (perhaps among others) needed to be triggered. First of all, a sentence initial pronoun has a referential function, picking out the current focal participant. My processor carried out that instruction correctly. Secondly, the pronoun's case marking sets up an expectation, either that this participant will fill the semantic role of *recipient* (e.g., an indirect object, as in *daj emu*, 'give to him'), or *experiencer* (as in *emu nravitsa*, 'he likes it', literally 'to him it is liked'). It is an empirical question which of these roles (discourse role or semantic role) a native Russian mental language processor would initially assign. However, in my case, we could argue that my Russian language processor did not carry out the second instruction at all, although it did carry out the first instruction. It identified the correct discourse participant (the speaker's son), but tried to relate him to a yet-to-be-encountered predicate in a way that ignored the case form.

It certainly does not appear to be the case that my Russian language processor *never* processes inflectional form to the level required in the above example. On the contrary, I would tend to feel that in many instances, my processor deals with inflectional form more completely. For example, I do at times find myself anticipating a verb which requires a "dative subject", based on the presence of a sentence initial dative noun phrase (i.e., in sentences such as *Emu nravitsa*, 'to him it is liked'). A variety of factors may conspire in the cases when I fail to implement specific details of processing required by the inflectional form of a word.

These introspective reflections may shed light on some of what was happening with participants in the experiments reported in Chapter 3. Those individuals may often have been hearing inflectional forms quite clearly, but at the speed of normal (if careful) speech, and especially when attention was partially drawn to the task of finding a picture which corresponded to the meaning of what was heard, inflectional forms remained inert.

My introspective experience is also consistent with the idea that when inflectional cues are not implemented, they quickly become unavailable for introspective observations, even if extra time is allowed for such retrospective observations. That is, by the time the whole sentence has been processed, it may no longer be possible to retrieve the particular inflectional forms of individual words from memory (although the lexemes can commonly be recalled in some form; typically the nominative form in the case of nouns).

It was suggested in the previous chapter that the Form-Oriented task depended to a significant extent on genuine processing mechanisms. In other words, the decreased processing demands of the Form-Oriented task did not merely increase the opportunities for metalinguistic analysis. In fact, I would suspect that in that task, too, clashing inflectional forms that went unprocessed online would often be unretrievable from auditory memory by the time a sentence had been fully comprehended, and thus there would simply be no possibility for participants to identify the errors with the aid of metalinguistic analysis. However, in that task there may have been an increased likelihood of fuller processing taking place online, thus leading to cue clashes. The argument to that effect in the previous chapter was based on the systematic nature of the reactions in the Form-Oriented task as opposed to those in the Printed Sentences task. The introspective examples discussed here could conceivably provide insights into why clearer effects of metalinguistic analysis would not have been observed in the so-called Form-Oriented task.

4.4. ON THE LENGTHY TIME-FRAME OF ACQUISITION

All else aside, one result of the present research appears secure. Having collected data from thirty-seven nonnative Russian users in Experiment 1, and fifty-four in Experiment 2, that is, ninety-one in the combined experiments, we are on fairly safe ground in saying that if a moderately nativelike system of inflectional processing does develop in adult learners of Russian as a second language, most features of the system, including what are pedagogically considered rather elementary features, do not develop quickly. Rather, the developmental process is spread over several years. Other aspects of the implications drawn from Experiments 1 and 2 may be more open to challenge. For example, the pattern of development sketched depended on categorizing error types generally on the basis of only two members of each category. To make a compelling case, it would be necessary to include a considerably larger number of examples of each error category, averaged out across a wider range of contexts. The reasonable systematicity of the findings and their theoretical plausibility encourage further investigation, but the interpretation of the findings provided here could well be overturned. On the other hand, it is extremely unlikely that we will find large numbers of nonnative speakers of Russian who have been learning Russian for four or five years, who have been living in Russia for two or three years, and who have developed moderately nativelike inflectional processing systems in relation to basic uses of case marking and aspect marking. That much appears secure.

The slow pace of the acquisition of the L2 Russian inflectional system must be in part influenced by the complexity of the system. In Chapter 3, I pointed out the many-to-many nature of inflectional form-to-function relationships in Russian. Slobin (1982) compares child acquisition of the ability to assign agenthood and patienthood in English, Spanish, Turkish and Serbo-Croatian. The comparison between Turkish and Serbo-Croatian is especially interesting in the present connection. At the age of 2;0 to 2;4 (=2 years;0 months to 2 years;4 months) the Turkish children had already achieved the level of ability to assign agenthood and patienthood that they and children of all other language groups showed by the age of 4;0-4;4. By contrast, the Serbo-Croatian children lagged behind the other groups until age 4;0-4;4, at which point they had nearly caught up. In Turkish, accusative case is marked by a single, consistent suffix (with systematic phonological alternations). Serbo-Croatian is like Russian (though perhaps not quite as extreme) in the many-to-many nature of case form-to-case-function relationships. If inflectional complexity prolongs the course of first language acquisition, then we should expect it also to prolong the course of second language acquisition insofar as the latter

depends on normal linguistic mechanisms. Furthermore, if children require four years or more to acquire the basic system of inflectional cues, it should not surprise us to find that adult learners take even longer, given the more limited role of Russian in their overall linguistic experience, possibly combined with biological disadvantages of adult language learners in comparison with child language learners.

Perhaps the complex nature or Russian morphosyntax helped to bring to the fore the limits of relatively small scale language learning experiences in relation to the genuinely linguistic developments. Of course, other L2s (more precisely, perhaps, other L1-L2 interlanguages) have their own "fossilizable linguistic phenomena". For L2 Russian, we are led toward a conclusion that exposure time required for relevant learning is not to be measured in scores of hours, or even in hundreds of hours, but rather in thousands of hours. The one person we discovered who was nativelike in her tendency to react to the inflectional errors of Experiment 2 is a person for whom Russian has been the constant primary language of everyday life for a matter of years. Is such a time-frame relevant only to L2 Russian? It may just happen that Russian provides us with a wealth of fossilizable features to investigate. There does not seem to be any obvious reason not to suspect that one crucial factor behind the fossilizability of any highly fossilizable feature in any language is inadequate experience in processing the language. No doubt this is far from being the only factor, but perhaps it is one that needs to be given more recognition. It seems likely that even with such massive experience with an L2, adult learners may still display fossilized features, or at least features with respect to which learning is proceeding so slowly that it appears to have ceased. However, it is really only fossilization in the context of massive experience that demands further explanation. In the absence of adequately massive experience with the L2 on the part of a learner, we might suspect that learning has not been complete because the minimum requirements have not been met.

To return to the level of personal experience, I find myself regularly encountering for the first time (as far as I know, judging by their unfamiliarity) forms of common words which I would guess would be familiar to preschool children. For example, I metalinguistically learned of the various forms of the verb *est'*, 'eat' at an early point in my Russian learning, but had been in Russia for over a year before I encountered the first person singular form, *em*, 'I eat' for what seemed like the first time. I happened to be spending a few days with a Russian family in a village at the time. I continue to regularly encounter first person and second person forms of verbs familiar to me in their third person forms. I feel that this reflects certain respects in which my input has been (perhaps dramatically) impoverished, both in quality and quantity.

For languages with less complex inflectional systems we may imagine that the time-frame of inflection-related acquisition will be shorter. However, other aspects of other languages, such as language-specific parsing strategies, may plausibly be expected to demand similarly large quantities of experience. For example, recall the much discussed L2 problem of the placement of English adverbs in relation transitive verbs (e.g., **I ate quickly my lunch*), and the fact that for native speakers, a postverbal adverb is a highly reliable cue that there will be no direct object in the clause (e.g., *I ate quickly and left*)? In what timeframe might nonnative English listeners acquire the parsing strategy exploiting that cue, and subsequently adapt their spoken production to the comply with that parsing strategy? One can imagine that such subtle parsing strategies might not be acquired quickly. This is, of course, an empirical question, but the difficulty researchers and teachers have had in effecting change in this area (Trahey & White, 1993) should lead us to suspect that researchers were looking for results to occur in an unrealistically brief time-frame (especially if *time* is understood in terms of total time of normal linguistic experience). In general, Selinker's mystery of fossilizable items discussed in the introduction may find much of its answer here. The acquisition of certain cue-process

sets occurs as a result of linguistic experiences far in excess of what the fossilized learners have been party to. This is imminently reasonable. Full-blown acquisition of language may not be in the cards when experience with the L2 is limited to what normally occurs in the workplace, in cross-linguistic contact situations, in a foreign language classroom, or in a typical immersion program. More appears to be needed. I suspect, much more. Perhaps research will eventually lead us to recognize that, from a neuropsychological perspective, languages are suited to complete acquisition when they play an overwhelming role in one's knowledge, experience and social life over the course of several years. Thus, adult learners may be growing in the right direction. However, if it takes a few years in the context of massive daily experience with a language to achieve full, healthy functioning, the more limited L2 experience of most adult learners may make the ultimate development of full, healthy functioning unlikely. There may commonly be no practical way around this limitation.

4.5. ON THE DEVELOPMENT OF PRODUCTION MECHANISMS

In Chapter 1 it was pointed out that in normal language development, comprehension mechanisms precede production mechanisms, so that it is not unreasonable to expect that in some sense, the comprehension system trains the production system. Zakharova (1973 [1958]) mentions a stage in preschool Russian language development during which children self-correct case forms that they utter, but she does not give any examples. Examples from English illustrating apparently the same general phenomenon can be observed in transcripts of the speech of the child Lise Menn studied and referred to as Seth (Peters & Menn, 1993). (These transcripts are available in the CHILDES database, available on-line from Carnegie-Mellon University.) One example occurred at the age of 2;11, when Seth was attempting to retrieve the word *stuck* and uttered the sequence *stu... sticked, stucked.* The other example occurred at the age of 4;1, when Seth went through the sequence: *taked, tooked, tooked.* In both instances, the forces driving speech production produced what is sometimes called an over-regularization. Something in Seth reacted to this output of his production system. In terms of the conception of learning developed in Chapter 1, what reacted in Seth would have been his comprehension system. The cohorts of forms activated by *sticked* and *taked* did not include forms compatible with the discourse context. The forms *stick, sticks, sticking*, and the forms *take, takes, taking* would have been in the cohorts activated. But none of these would be associated via the comprehension lexicon with foreground events of taking or sticking. The next attempts, *stucked*, or *tooked* would fare better in that at least they would activate the target word-form in the comprehension lexicon, since the target form, *stuck* or *took* is embedded in the form actually produced. Based on her remarks, it appears that Zakharova observed analogous cases of Russian children tuning their case forms. I would claim that such tuning gives us a window into the interaction of the comprehension system and the production system, as the former trains the latter.

There might appear to be a problem with this conception of learning in the case of adult L2 Russian learners at the levels of development represented by the experimental participants in Chapter 3. Many of them would probably intuitively feel that, when it comes to the elementary uses of cases, aspects and agreement that were involved in the experimental stimuli, their spoken production would have been more nativelike than the corresponding comprehension mechanisms might allow for, insofar as those comprehension mechanisms are tapped by Experiments 1 and 2. That is certainly the case for my own spoken Russian. In other words, it might well seem to be the case for many that production ability has developed ahead of comprehension ability. To claim this with any confidence, we would need adequate samples of spoken production and on-line

comprehension data from the same participants. However, suppose we accept this intuitive impression as valid—many L2 Russian learners are more consistently nativelike in certain features of their production than in the corresponding comprehension processes. Where could that production ability come from, if the comprehension system is not in a position to tune the production system? To answer that question we need to consider the possible independent contributions of normal linguistic mechanisms on the one hand and other speaking strategies involving metalinguistic planning, on the other hand. Keep in mind that by *metalinguistic planning* here is generally meant the application of so-called explicit grammatical knowledge or observations.

Suppose we are correct in attributing the advantage of the so-called Form-Oriented task over the Meaning-Oriented task to different levels of processing demands. Then we might suspect that relevant inflectional processing events are variably present, more likely to be present when overall processing demands are low than when they are high (VanPatten, 1996). Only when reactions to inflectional cues become consistently strong could we expect the comprehension system to bring the production system fully into line. That is, until that time, the production system can continue to produce non-target-like forms that go unnoticed by the comprehension system. It may take several years before the learners comprehension would provide strong, consistent reactions to some particular variety of nonnative output feature. As the comprehension system gradually strengthens with regard to some feature, the production system might become more and more consistent as well. Variability would be inherent in the learning apparatus. Without some additional influences being brought to bear, production mechanisms might never be more consistently nativelike than comprehension mechanisms. In the common case where the nonnative speaker never does come to have adequate experience for full nativelike functioning of the comprehension system, the spoken production system would never become fully nativelike either.

As noted, it does often appear that some nativelike production patterns can develop too far ahead of the corresponding comprehension mechanisms to account for them by appealing to the variable nativelikeness of the comprehension mechanisms. Nativelike suppliance of oblique case forms can occur very early, with almost no time lag between initial pedagogical exposure to them and their use in some spoken or written productions. Such production must be largely independent of the comprehension mechanisms which process inflection. That is, if a learner as yet has almost no tendency to react to particular inflectional cue clashes, then it would seem that s/he would not detect errors in her own speech, except via metalinguistic analysis. This is evidently the case. In classes that I attended in Russia, I commonly observed learners (including myself) who uttered nonnativelike forms and showed no awareness of them. Yet when corrected, the L2 speakers could demonstrate the relevant metalinguistic understanding of their errors, for example by naming the case value or aspect value that they should have used and giving some indication of why. When such learners do self-correct, it could be difficult to tease apart the involvement of metalinguistic monitoring and inflectional processing.

It could be argued that the results in the Printed Sentences task provide key evidence in this connection. In the listening tasks of Experiments 1 and 2, the auditory word forms flowed by and were gone, while processing had to go on. I suggested, on the basis of my own introspective observations, that the word forms often rapidly cease to be retrievable from memory in a detailed auditory format. The difference between the listening tasks and the Printed Sentences task appears to indicate the same thing. There was considerable improvement in the Printed Sentences task, where the written word forms are perceptually stable, giving the reader control and flexibility in repeatedly processing and analyzing the sentences. The reader is free to backtrack as many times as s/he wishes. S/he can apply metalinguistic knowledge to a tentative revision, check the

revision out mentally by rereading the sentence with the revision supplied and so on. This task thus has more in common with spoken production than do the two listening tasks. If the listener can apply metalinguistic knowledge in order to supply the desired word-forms in the Printed Sentences task, then we can imagine that s/he can often do the same when deliberately planning spoken production. Unlike listening comprehension, the content and form of spoken production is somewhat under the speaker's control and s/he can wait and speak when s/he feels ready.

Some studies have in fact demonstrated that allowing extra time for a production task does increase the incidence of target-like use of some grammatical morphemes. Ellis (1987) found that L2 use of English regular past tense forms in required contexts was greatest in a written composition task, least in an unplanned oral production task, and intermediate in a planned oral production task. In the composition task and the planned oral production task, the same narratives were used. Thus many of the same verbs occurred in the same contexts. Ellis points to one subject who supplied the past tense forms of verbs in the first half of the planned oral production task, but appeared to tire out and revert to base verb forms without past tense marking in the second half of the narrative. These were verbs that had been written in the past forms in the same portion of the narrative in the composition task. Ellis' third task, which required the production of a narrative after only enough time to grasp the events themselves, showed a substantial further decrease in the ability to produce the past tense forms.

Crookes (1989) likewise found that L2 use of the English definite article became more target-like under a planned production condition than under an unplanned production condition. However, effects of planning for other grammatical morphemes (plural marking, indefinite article, plural agreement) were either nonexistent or non-significant. Without access to his qualitative findings, it is difficult to interpret the difference between definite articles and other grammatical morphemes.

Foster and Skehan (1996) also compared unplanned production with two kinds of planned production. Their "accuracy" variable was unaffected by planning condition. Since the measure used was the percentage of error free utterances, however, we cannot know what an analysis focusing on grammatical morphology might have yielded in terms of evidence for effects of planning. However, in measuring "syntactic variety", Foster and Skehan did find an increased use of "nonsimple past tenses" in the narrative task under the planning conditions in comparison with the unplanned conditions.

Except for the composition task used by Ellis (1987), these studies did not allow scope for utterance-by-utterance planning. Once a task was underway, participants attempted to sustain continuous production. We might expect that a larger effect of planning on target-like use of grammatical morphology would occur if metalinguistic planning were allowed on an utterance-by-utterance basis. It could be argued that in effect, this is what was observed in Ellis' composition task. The target-likeness of the English in the composition task in contrast with the English in Ellis' spoken production tasks is striking.

An important study by Hulstijn and Hulstijn (1984) showed no effect of time pressure on the production of two target structures. However, the task involved retelling short passages rather than totally free production. It is possible that without the immediately preceding native model productions, the participants would have been less nativelike in the production of the target structures under time pressure than without time pressure, since that would have made metalinguistic planning more demanding.

The ability of people to reflect metalinguistically seems uncontroversial. As Felix (1987) notes,

> A linguist analyzing an unknown language on the basis of oral or written material, a student attempting to write a grammar for a given set of data, or a language teacher pondering over the most effective way to present language data and rules to his students, are typical cases: they appear to process language data with essentially the same system of cognitive potentials that they utilize when trying to gain knowledge in any other intellectual domain. (Felix, 1987, p. 156)

If this is correct, then we might also expect that the application of such knowledge would change in certain respects over time. Initially production might be slow and perhaps verbally mediated (as when a learner verbalizes silently to himself or herself that a certain word represents an instrument and therefore must be in the instrumental case). Eventually the planning might take place more quickly, with less awareness (as in the learning process described by Anderson, 1982, in connection with high school geometry).[24]

Felix (1987) argues that metalinguistic thinking about the L2 is virtually unavoidable for people who are at a cognitive maturational level where such thinking is possible. This is powerfully illustrated by a diary entry in my possession from a thirteen-year-old boy who was asked to write about his experience and feelings following language learning activities. In the language learning session he describes, the learners performed a task that involved picking up vinyl stickers in response to instructions given by a native speaker. They had been learning Russian in Canada several hours per week for two and one half months. They were familiar with the nouns referring to the objects depicted by the stickers, but mainly in the nominative and accusative singular forms. In the activity the Russian speaker would instruct the learners to pick up either singular objects of plural objects: *Pick up the boy* (*mal'chik-a*, 'boy-**accusative:singular**'), or *Pick up the boys* (*mal'chik-ov* 'boy-**accusative.plural**'). The stickers were organized so that there was an individual example of each (e.g., a lone boy sticker) and a collective set of each (e.g., a group of three or four boy stickers). After a few minutes, the learners were able to rapidly and correctly respond to the Russian instructions, making use of the number inflection as required for comprehension. The task was similar in spirit to VanPatten's instructed input tasks (VanPatten & Cadierno, 1993; VanPatten and Sanz, 1995; VanPatten, 1996). However, no explicit instructions were considered necessary, as the task itself was initially demonstrated so as to highlight the difference between singular and plural forms. As the planner of the task, my goal was for the learners to form direct associations between the plural morphology and the plural meaning, avoiding the need for any metalinguistic reflection. That is, the plural forms would be made immediately meaningful, the stage of grammatical explanation thus being made unnecessary. The diary excerpt below implies that the terms *plural* and *ending* must have come up in the discussion surrounding the activity. However, there was no metalinguistic presentation or discussion of the formal differences between singular and plural nouns (or nominative-accusative differences). In spite of the absence of metalinguistic explanations, the following diary entry (used with permission) illustrates a high level of metalinguistic analytical thinking on the part of a thirteen-year-old boy:

---

[24] Of course, the danger in discussing such learning in connection with second language learning is that it is an aspect of learning that is highly apparent to learners: early L2 use is extremely effortful and provokes metalinguistic planning, while later L2 use is considerably easier. Unfortunately, the more purely linguistic aspects of learning are not apparent to learners, and thus may be underrated in popular thinking.

> We just finished our lesson with M and it was interesting. We did some plurals with the stickers. I noticed that with almost all the things the endings were the same. When it was a plural there was an e sound at the end. When it was pick up the plurals the vowel was dropped off completely. When it was singular it was just the word but pick up the singular for words without a vowel ending it added a vowel sound a. if it already had the vowel ending a then it changed to o. Other than that I didn't really notice much else.

Most of this should be intelligible to an instructed learner of Russian (the boy heard the ending –u as "o"; he transcribes [i] and [ɨ] both as "e" and [a] as "a"). This diary entry supports Felix' claim. Adolescents and adults (if not younger children) are at least capable of such metalinguistic analytical thinking. Individuals may differ considerably in their ability to make such generalizations. One of the constructs proposed in the Modern Language Aptitude Test (MLAT—Carroll & Sapon, 1959; see Ehrman 1996 and Skehan, 1998 for recent discussions), *Words in Sentences*, is claimed to be a measure of sensitivity to grammatical patterns. The ability or abilities tapped in that section of the MLAT would presumably influence the ability of a given individual to make analytical metalinguistic observations, if not the ability to gain and retain such knowledge in the first place.

It appears likely then, that to varying degrees with different individuals and under different communication conditions, L2 speakers might be able to exploit metalinguistic knowledge in planning speech, and in so doing, be able to make their spoken production appear more nativelike than it would on the basis of the purely linguistic mechanisms at work within the speakers. The idea of two types of cognitive processes—linguistic and nonlinguistic—playing a role in L2 production has been frequently proposed (Krashen, 1977; Felix, 1987; Bialystok, 1978; Schwartz, 1993; Paradis, 1994; Sharwood-Smith, 1996; Truscott, 1998;). This possibility raises enormously interesting questions that are rarely (if ever) addressed in the literature on metalinguistic awareness and/or grammaticality judgments (e.g., Chaudron, 1983; Bialystok & Ryan, 1985; Masny & d'Anglejan, 1985; Birdsong, 1989; Gass, 1994; Sorace, 1996). Rather, the objects of metalinguistic awareness that are believed to be available to the everyday intuitions of literate language users (including these writers themselves) are largely taken for granted. In actual fact, many aspects of normal language processing outlined in Chapters 1 and 2 are not obviously available in the form of subjectively identifiable experiences. (If they were, then we could simply ask L2 learners whether, for example, in using perfective aspect they were treating it as a marker of inherent verbal aspect, or as an instruction to increment the mental model by one event). As discussed in Chapter 1, the units of observation which are at times taken for granted in discussions of metalinguistic strategies—for example words, syllables, segments, sentences, and membership of words in grammatical categories—are not present in the speech stream, but are rather the result of processes in the listener which differ in their details form language to language. There is a fundamental question of what makes certain results of processing, say phonological words, or the grammatical category noun, available for objectification and discussion.[25]

---

[25] In a modular approach to language processing, it might be suggested that the points at which language processes are accessible for introspective awareness are the points at which separate modules exchange information (see Friederici, 1990). However, to account for the objects of metalinguistic awareness in such way might be too restrictive. Under the right circumstances, people learn to identify language processing experiences that were previously not objectified within their speech communities. Examples include words, syllables, vowels and consonants, but also the more esoteric categories

Suppose people can come to recognize the grammatical category *noun* because all words belonging to that category trigger analogous processing instructions, i.e., instructions to add tokens to mental models, for example. Mental language processors learn to expect such instructions in certain contexts, such as following words that trigger other kinds of processing instructions (transitive verbs, for example). This would imply that the relevant processing experiences must develop (to the extent that they are not biologically hard-wired) before the objects of metalinguistic observations can be available for detection and categorization.

For the native language, this conclusion may not be problematic, since not much metalinguistic observation goes on before extensive processing ability is in place. For an early second language user, as for a field linguist analyzing data elicited from informants (or linguistics students doing a morphology assignment), the objects of metalinguistic observation must be of a different nature than the objects of metalinguistic observation in one's native language, since the relevant internal processing experiences do not yet exist. Also, translation into the learner's or analyst's L1 would appear to play a crucial role, in such cases. In other words, although the term *metalinguistic* is commonly applied indifferently to L1s and L2s, in fact, the type of metalinguistic grammatical knowledge that might allow an L2 user to mimic nativelike production patterns in advance of the development of their comprehension functions must be a fundamentally different phenomenon from L1 (and truly nativelike L2) metalinguistic knowledge. L1 metalinguistic knowledge involves intersubjective identification of shared language processing experiences. Early L2 metalinguistic "knowledge of rules" is on a different plane, involving the identification of L2 patterns (typically in written text), and perhaps involving strategies for equating those patterns with L1 processing experiences, since those patterns have little or no relationship to the L2 users own L2 processing experiences. On the other hand, as the relevant L2 processes come into existence, they create the possibility of genuine, L1-like metalinguistic awareness.

These two different varieties of metalinguistic awareness may have been evidenced by the fact that in the Printed Sentence tasks in Chapter 3, for many learners it appeared that successful acquisition and use of L2 metalinguistic ability depended in part on increased proficiency. That is, it may be that many learners have trouble acquiring the sort of early L2 metalinguistic knowledge that has no basis in the learner's own language processing experience, A few intermediate learners did score high in the Printed Sentences task, while showing almost no sensitivity to inflection, as indicated by reactions to cue clashes, in the listening tasks. For the most part, however, it would appear that the metalinguistic knowledge concerning the most elementary uses of the cases, covered in the first few months of formal instruction, becomes significantly more secure after a few years of L2 development. That is, apparent learner gains in the context of a controlled classroom activity might not represent secure gains, even in the acquisition of metalinguistic understanding. For many learners, it would appear that the availability of metalinguistic knowledge that might help them to simulate nativelike production is limited by the state of development of their L2 language processing systems. Thus on the one hand, for a given learner at a given point in time, the current levels of metalinguistic analytical understanding might enable some simulation of nativelikeness before the relevant nativelike linguistic mechanisms have taken hold or matured. On the other hand,

---

identified and discussed by specialists, e.g. allophones and scope ambiguities. The particular set of objectified linguistic experiences in a given culture would seem to be a product of the history of that culture, but there may be a wide range of possible intersubjective identifications and categorizations of aspects of linguistic experience.

for many learners, the maturity of the L2 mechanisms might also place a limit on the level of analytical understanding that is attainable.[26]

We have already noted the possibility of metalinguistic production strategies themselves would become increasingly efficient over time. Anderson's ACT* model of skill learning (Anderson, 1982, 1983) includes a process called *composition*. In composition, several steps involved in achieving a cognitive goal come to be replaced by a single step. To take a trivial example, an individual might work out the sum of seven and three using a counting strategy the first few times that the sum is needed, and then subsequently retrieve the sum from memory whenever it is needed. Similarly, in the application of metalinguistic knowledge, we might expect that a particular problem would not be need to solved through the same steps after the problem had already been solved several times. For example, with common direct objects of a given verb, the L2 user might use a verbally characterizable (and perhaps verbally mediated) metalinguistic strategy to produce the accusative form. S/he might reason that since *pivo*, 'beer', pronounced ['pivə] is spelled with an *o* in Russian orthography, the form will not change to *pivu* when it is the object of *pit'*, 'drink', as it would if it were spelled with a final *a* (as in the case of *vodka*, 'vodka', pronounced ['vodkə]). But once the speaker has uttered forms of *pit'*, 'drink' followed by the noun *pivo*, 'beer' (or *vodku*, 'vodka:accusative') several times, s/he could, by the principle of compilation, simply retrieve the phrase *pit' pivo*, 'drink beer' (or *pit' vodku*, 'drink vodka) from memory fully formed.[27]

The ability to reuse earlier productions (presumably with increasing ease) has been demonstrated in recent studies of task repetion. Bygate (1996, cited in Gass, et al., 1999) had learners narrate a video segment twice, with a ten week interval in between performances. He noted the reuse of specific phrases from the first performance during the second. Gass, Mackey, Alvarez-Torres and Fernández-Garcia (1999) examined improvments in the use of the Spanish copula *estar* (an item which typically appears to be difficult for L2 Spanish learners to acquire) and found improvement when learners repeatedly narrated the same video segment, but not when they narrated different video segments. Importantly, this did not represent learning that then carried over to a novel video. If this limited finding were to prove more general, we might expect that L2 users would appear more nativelike when discussing topics they have frequently discussed previously than when discussing topics for the first time, based possibly on their ability to reuse production solutions from earlier communication efforts. In practical terms, pedagogical error treatments which result in successful learner-generated reformulations

---

[26] Green & Hecht (1992) found that adolescent L1 German learners of L2 English were only able to demonstrate analytical grammatical knowledge of about half of the English errors that they were able to correct (errors that had been self-corrected by similar L2 English users in the study reported in Green & Hecht, 1993). That is, their performance in a pencil-and-paper error correction task could only be partly attributable to the type of metalinguistic understanding taken to be a major factor in our Printed Sentences task. As we do not know how the same participants might have performed in an online listening task, it is hard to compare those results with the current results. It is possible that the ability to "give a correct rule" would correspond to increased error detections in a pencil-and-paper task vis-à-vis listening tasks. The detected errors for which no metalinguistic analysis was available to the learner may be the subset that reflect truly linguistic mechanisms. If so, then Green and Hecht's results might be interpretable in a manner comparable to the interpretation of the results reported in the previous chapter of this dissertation, where the unrestrained application of metalinguistic knowledge appeared to lead to an increase in error detections

[27] It is important to distinguish between formulaic language, where the speaker may have little sensitivity to the potential internal constituents of utterances, and prefabricated utterances which have can be retrieved ready-made on the basis of their previous productive use in production or comprehension.

(Lyster & Ranta, 1998, Lyster 1999) could benefit L2 production in terms of adding to the stock or ready-made verbalizations without having a comparable impact on the relevant form-function associations within the language processing system.

Another source of ready-made utterances and utterance components might be the storage, not of learner-generated utterances, but of native utterances which the learner has heard, and particularly those which s/he used in her own production shortly after hearing them. Such an implication might be drawn from Hulstijn & Hulstijn (1984), where, when instructed to do so, learners were able to recall and produce specific grammatical patterns in a passage-retelling task.

Skehan (1998, drawing on Bolinger, 1975, and other research) attributes a major role to the deployment of such already familiar phrases during speech production, although not distinguishing between ready-made phrases which might result from normal language production processes, those that might result from metalinguistic problem-solving strategies (a difference being that in the case of the latter inflectional form might remain inert to the comprehension system, while being active in the case of the former), and those that involve the adoption of native-produced utterances. More broadly, a learner might carry out a large amount of metalinguistic problem-solving when discussing a particular topic in the target language for the first time, since a new topic might require many new lexico-grammatical collocations. Once the learner has discussed the topic many times, his or her speech may have many features that appear to be nativelike as a result of compilation or some such process. In addition, phrases that are familiar from frequent encounters in comprehension could add to the stock of retrievable native-sounding (we might say, *pseudonative*) productions.

In recognizing the possibility of various types of precompilation of native-like patterns and speeded up metalinguistic strategies, we should perhaps be cautious in attributing too much credit to such nonlinguistic mechanisms in the spoken production of near-native L2 users. It is possible that near-nativelikeness is only possible as the result of a reasonable amount of genuinely linguistic learning dependent upon many thousands of hours of experience with a language, both as a listener and as a speaker. At that point, the role of nonlinguistic mechanisms might be greatly diminished.

4.6. SUMMARY: SOURCES OF PERFORMANCE

At this point, it might be helpful to summarize the variety of skills that might distinguish the listening tasks of Chapter 3 from the Printed Sentences task, and that might distinguish between the relatively more native-appearing spoken productions (assumed, but not specifically investigated in this study) and the relatively less nativelike comprehension mechanisms (and by implication, form-function relationships).[28]

[28] Initial L2 learning inevitably involves changes in some initial state. From a language processing perspective, there would seem to be a natural disposition toward a "full transfer" position (Schwartz & Sprouse, 1996). That is, during the first moments of attempted L2 processing, the L1 processing system must do whatever it can to cope with the input. Later, if it is true that contrasting grammatical forms and patterns are often initially treated as exemplars of a common prototype (the contrast being neutralized), then the learner's L1 might influence his or her choice of the prototype, which might in turn influence the patterns of production. The idea that the initial L2 processor is the L1 processor is also shared by MacWhinney (1997) and VanPatten (2000). This primarily addresses "transfer in comprehension" (Ringbom, 1992). The attempted employment of L1 production strategies for L2 production may be a more tumultuous affair, due to the lack of correspondences between L2 and L1 resources.

1) Normal linguistic learning: As cue-process pairings occur in working memory cues become associated with processes in the comprehension system, so that the cues are increasingly able to trigger the processes. This type of learning affects performance (it is claimed) in the listening tasks of the previous chapter, even when there is little opportunity to employ metalinguistic problem-solving strategies.

2) L2 metalinguistic learning: Learners differ in their ability to acquire metalinguistic knowledge of a purely analytical nature, when there are still no learner-internal processing experiences to relate that knowledge to. To the extent that learners do acquire such knowledge ahead of the corresponding linguistic processes, it provides the basis for some of the improved performance demonstrated in going from the listening tasks to the Printed Sentences task. For many learners, improvements in natural linguistic mechanisms allow an increase in L1-like metalinguistic knowledge (that is, related to internal processes which the language user actually experiences). As the level of this latter type of metalinguistic knowledge increases, it opens the way for new learning of the more purely analytic (instructed L2-style) metalinguistic knowledge.

3) Metalinguistic strategy learning: As learners verbalize message employing metalinguistic knowledge, analogous verbalization problems repeatedly arise. Over time, this leads to increased ease in the use of such strategies (whether or not it leads to true automatization).

4) Adding to the stock of ready-made expressions: In verbalizing messages, L2 speakers can reuse words, phrases, and whole utterances with increasing ease as the same or similar messages or message components are repeatedly reverbalized.

Achieving near-native proficiency (that is, native-like form-function relationships and expectations in comprehension, which can in turn guide production) may require a large amount of the first variety of learning. With respect to some aspects of target languages, for example, Russian inflectional morphology, this may require massive experience with the language over the course of a number of years. Meanwhile, the other types of learning might enable the simulation of many nativelike features of production.

Studies demonstrating effects of pedagogical interventions might be detecting any of these four varieties of learning, or some combination of them.[29]

4.7. THE LONG ROAD AND MODELS OF SLA

The findings of the previous chapter appear to be compatible with a picture of SLA which includes a gradual accumulation of strength of processing cues for purposes of comprehension, which in turn, given certain not unreasonable assumptions, predicts gradual development of corresponding production patterns. In the case of production, the natural patterns of strengthening might be masked to a small or large extent by the presence of nonlinguistic cognitive strategies which can mimic aspects of nativelike production. However, not all aspects of language learning are as prolonged as the development of inflectional processing in L2 Russian. Lexical learning may occur relatively rapidly and may carry with it the acquisition of semantic argument requirements

[29] This proposed set of learning categories has something in common with the model in Towell & Hawkins (1994) in that it combines normal linguistic learning (in their case, UG-based learning) with other kinds of learning. Unlike Felix's (1987) proposal, linguistic and metalinguistic learning are not seen as being in competition. On the other hand, unlike information processing approaches (e.g., Skehan, 1998), I would not pin an excessive amount of hope on the prospect of nonlinguistic mechanisms doing the lion's share of the work over the long haul. Recall from Chapter 1 that L2 systems are faced with the need to solve the same basic computational problem that L1 systems solve, and would appear to be largely dependent on the same physical equipment at many levels.

of verbs and other selectional and collocational properties. Basic word order cues (such as those involved in distinguishing grammatical subjects and direct objects) may be acquired comparatively readily, while others (meaningful permutations, such as those involved in question formation in English or subordination in German) may present a picture of long-term, gradual learning. It seems apparent in the light of the experimental results reported in the previous chapter, that the acquisition of case, aspect and verbal gender agreement in Russian are subject to prolonged development.

4.7.1. SLA as the Acquisition of Competence Grammars

Insofar as models of SLA draw on linguistic theories, it seems likely that they will not easily accommodate the notion of gradual learning (as opposed to categorial learning such as rule acquisition, parameter setting, the addition of grammatical patterns to a mental inventory of permissible ones, etc.).

The notion of modality-neutral linguistic knowledge has perhaps been most clearly articulated by Chomsky, in terms of competence (Chomsky, 1965) and I-Language (Chomsky, 1986, 1995), which is understood as "a generative procedure... that generates *structural descriptions*... each a complex of properties, including those commonly called 'semantic' and 'phonetic'" (Chomsky, 1995, p. 167, italics original). It is hard to see a role for gradual learning within a competence grammar. Although the existence of learning within the performance systems may not be ruled out, Gregg (1996) implies that learning within the performance systems would be outside the domain of SLA:

> The domain of a L2 acquisition theory is not the behavior of speakers (linguistic performance), but rather the mental system (competence) underlying that behavior. (p. 53)

This excludes the possibility that much or most of the learning that goes into SLA is actually located in performance systems. For example, the direction of thinking Gregg is expressing could exclude the originally central question of fossilization from the realm of SLA questions.

Some non-Chomskyan approaches also treat grammar as an internal specification of a language, or as an inventory of patterns or an inventory of form-function relationships. Langacker for example takes grammar to be "not... a generative or constructive device, but simply an inventory of conventional linguistic units" of various sizes and various levels of specificity vs. schematicity (Langacker, 1991, pp. 227-28). Rather than sentences being generated or not generated by the grammar, they are sanctioned or not sanctioned by its schemas.

It is less clear what exactly would be acquired under a typological approach to grammar, as typological approaches are typically concerned with concrete samples of speech as produced in real speech situations, rather than with internal grammars (Comrie, 1981). However, Ferguson (1987), writing in the general Greenbergian tradition in connection with SLA, makes reference to "repertoires of structural units at many levels", saying that "languages differ greatly in the size and composition" of these repertoires (Ferguson, 1987, p. 247). Similarly, Eckman (1996) talks of acquisition in terms of the presence of structures in languages (as when the presence of one structure implicationally predicts or favours the presence of another). In general then, linguistic theories precondition us to view the "what" of language learning in terms of the patterns observed in speech, or determined by introspective reflection and thus to view learning a language as the acquisition of these patterns, or of "knowledge" of these patterns.

4.7.2. The Monitor Model

Turning to SLA models as such, one of the most influential models has been Krashen's Monitor Model (Krashen, 1977, 1985, 1981, 1987), although it no longer has wide support (if it ever did—see virtually any introduction to SLA for a summary of the standard objections, e.g., Archibald & Libben, 1995; Mitchell & Myles, 1998). Central to that model is the Input Hypothesis, which holds that the internal grammar is constructed by a learning device on the basis of comprehensible input. From the perspective of this dissertation, learning to deal with the input in the first place constitutes a large part of the learning that is of interest (see also Carroll, 1999). In Krashen's approach there is no emphasis on learning to process the input. Rather, it appears that the ability to process the input is in place from the start and feeds the system that organizes a grammar based on evidence in the input. Thus, neither is there any place in for the learning involved in the development of production mechanisms. Neither the development of comprehension mechanisms and nor the development of production mechanisms has a place in Krashen's model.

4.7.3. Gass' Model of Input and Interaction

A recent approach which appears to have somewhat of a Krashenian heritage is that of Gass (1997). Gass' model, like most, starts with input and ends with output. Output has various roles in her model, such as fostering the negotiation of meaning and hence generating useful input, as well as testing hypotheses, thus directly affecting the learning mechanisms which produce the internal grammar. Assuming that Gass does not in fact believe that the internal grammar has no role in the comprehension system, we might imagine a modification of her model such that the final stage is not *output*, but rather *performance* more generally, including both comprehension and production. Like Krashen and Gregg (and UG-oriented researchers in general), she does not strongly emphasize the acquisition of L2 comprehension mechanisms or production mechanisms as such. Provide the right input under the right conditions, combine it with production activities and the grammar develops. Develop the right grammar, and comprehension and production mechanisms will be there. However, the experimental evidence presented in Chapter 3, consistent with my introspective observations reported in this chapter, suggest that the readiness with which details of input form get processed depends in part on the extent to which they have already begun to take on processing functions. This would appear to involve a spiral effect in which input gets processed increasingly as the form-function relationship gets stronger. Form-function relationships might first gain a small foothold, after which they would need to grow in strength until they function consistently and robustly. In Gass' model however, as in Krashen's, it appears that the forms are simply there in the input, are apperceived (that is, recognized as something needing to be learned) and then exploited by the learning mechanisms which construct a grammar in which these forms immediately have categorical status.

4.7.4. VanPatten's Input Processing Model

Of the models considered here, VanPatten's (1996; 2000) input processing model may have the greatest number of points of contact with the conceptions developed in this dissertation. Like Gass' model, it begins with input. However, rather than a grammar construction process which results finally in *output*, learning leads to what VanPatten calls the *developing system*. The slow pace of development in relationship to grammatical morphology in general is attributed to what he calls the low level of communicative value

of grammatical morphemes. As long as processing resources are exhausted dealing with the content words, there is insufficient capacity left for detecting grammatical features and processing them for purposes of acquisition. However, various factors, such as pauses in the input, might decrease the demands placed on processing resources. This is in line with the improved performance we saw in Chapter 3 in the Form-Oriented task in comparison with the Meaning-Oriented task. There, it was suggested that the decreased processing demands of the Form-Oriented task enabled more inflectional processing, creating more opportunity for participants to experience cue-clashes. The instructions for participants to pay attention to form, I argued, may have had a smaller effect in the increased level of error detections. In VanPatten's model, deliberate attention to specific aspects of form can enhance inflectional processing, if it enables the learners to derive meaning from the form that they might otherwise have ignored. On the other hand, VanPatten (1990) showed that purely metalinguistic monitoring for form lowers the level comprehension of achieved by learners.[30]

VanPatten, as noted, proposes that the tendency of linguistic elements to be acquired earlier or later is related to their *communicative value*, which is positively affected by *inherent semantic value* and negatively affected by *redundancy*. On an intuitive level, this proposal appears to have much going for it. Utterances can often be correctly understood with grammatical elements omitted, while they cannot generally be understood with the lexical elements omitted. However, it may be that the concept of communicative value will be found to be secondary to other notions which are more explanatory. VanPatten proposes a learning principle according to which "learners process input for meaning before they process it for form"[31].

It might be argued that the principles of allocation of attention by the processor ought to fall out from deeper facts about the processing mechanisms. For example, associations between time-stable concepts (expressed by content words) and phonetic form take hold relatively easily. The processor may then make use of the lexical contribution of an inflected word-form for the simple reason that it is able to do so. At that stage, if the processor does not exploit inflectional details, it is simply because it has no use for them. Functionally, it is as though they aren't there. It will take a long time for that situation to change significantly, due to the brief nature of the processing events that inflectional forms must link to, and the fact that lexical processing is going on more or less simultaneously, exacting a processing cost.

Chapter 2 of this dissertation was intended to redeem grammatical morphemes, and inflectional morphology in particular, from relegation to the status of redundant, not very useful, formal features, accidents of linguistic history and perhaps more of a

---

[30] VanPatten's (1990) form-focused tasks involved listening for meaning while simultaneously non-meaningfully monitoring for a particular grammatical morpheme. These tasks are different in nature from the tasks reported in Chapter 3 of this dissertation, where there was no monitoring for particular morphemes, but rather reactions to errors.

[31] VanPatten (1994) proposes a set of principles related to this claim; briefly, these include the following: 1) learners process content words first; 2) they process lexical items in preference over grammatical items; 3) processing of non-meaningful form becomes possible once the more meaningful content can be processed with low attentional demands; 4) sentence-initial nouns are assigned agent status unless semantic considerations override; 5) only after grammatical cues are incorporated into the L2 system is the preceding principle relaxed; and 6) learners process sentence-initial elements first, sentence-final elements second, and sentence-medial elements last. The first three of these principles are the ones most related to the issue of inflectional morphology. The principles as stated do not indicate how the learner determines what are the content words and what are the function words, or what are the lexical elements and what are grammatical elements, or which element is relatively more meaningful than which other element.

nuisance than a help. It was argued that they play a lively role in normal L1 comprehension. Their importance is further reflected by the fact that languages are continuously developing new grammatical morphemes as old ones erode. They have a different function than content words, but from the standpoint of language processing at normal speeds, they clearly play a crucial role. If they are missed in processing, some of the work they do may have to be done inferentially. In the case of a full-blown native language processing system, this could in fact make speech that is lacking in the obligatory grammatical morphemes more demanding to process than speech in which they are supplied.[32]

VanPatten argues that there are differences of communicative value even among grammatical morphemes and that these differences may predict their ease of acquisition relative to one another. This view is supported by Kempe and MacWhinney's (1998) evidence that the increased cue-validity of case-marking in Russian as opposed to German leads to earlier exploitation of case-marking in L2 Russian than L2 German. However, there would appear to be aspects of grammatical form with intuitively very high communicative value which are nevertheless not readily acquired. An example from English might be the preposing of auxiliaries in question formation. A sentence initial auxiliary is a highly valid cue as to the (direct) illocutionary force of a sentence, and the difference between questions and statements is of considerable communicative import. Yet this grammatical feature is not acquired particularly readily (Larsen-Freeman & Long, 1991, Chapter 4; Pienemann, 1998). Johnson and Newport's (1989) findings indicate that adult learners have a comparable levels of difficulty with English particle movement (*He put out the candle* vs. *he put the candle out*) and English yes-no question formation. Yet they would seem to differ strongly in meaningfulness.[33] An alternative explanation of the difficulty of English question formation is that interrogativity does not involve a concept derived from perceptual experience having time-stability like that of concrete content-word concepts. Therefore, it could be argued, there is less opportunity for form-meaning associations to take hold on a given occasion in the case of interrogative marking than in the case of a concrete noun, verb, or adjective.

In any case, it has been argued on various grounds that content words and function words belong to separate processing subsystems in native speakers (Friederici, 1985, 1990). The difference in ease of acquisition of lexical morphemes as compared with grammatical morphemes may reside in their difference in function (and hence in the nature of the processes they trigger, and the relationship of cues to processes). If this turns out to be correct, then it may not be helpful to attempt to characterize the difference between content words and function words on a quasi-quantitative dimension of a differentially shared property of meaningfulness.

VanPatten argues that the relative ease of processing of lexical content (and the use of lexical content in sentence comprehension) determines the likelihood of grammatical material being processed. The difference in performance between the two

---

[32] A question might be raised regarding whether the fact that grammatical morphology is such a powerful force in L1 processing implies that it will also have an important role in L2 processing. Keep in mind that it takes a few years for the functions of L1 grammatical morphology to fully develop. That is, VanPatten's principles are probably as true of L1 acquisition as they are of L2 acquisition. Whether grammatical morphology ever becomes a powerful force in L2 processing is an empirical question. If it does not, then the difficulty in the development of form-function associations for grammatical morphemes might be an alternative explanation to an explanation in terms of their place on hypothetical scales of meaningfulness and redundancy.

[33] It may be that the availability of intonation as an interrogative marker negatively influences the acquisition of subject-auxiliary inversion as an interrogative marker.

listening tasks in the experiments of Chapter 3 may have been in part due to the fact that the increased processing demands in the first task (concurrent picture selection and shorter interstimulus interval) decreased the ability of listeners to process inflectional form. This would be in line with VanPatten's proposal. A related point was made in Thomson (1996), where it was hypothesized that, as particular layers of L2 processing become automated, new layers of linguistic elements might naturally rise to the notice of the processor. This could imply a modification of the prediction that acquisition of nativelike or near-nativelike inflectional functioning in L2 Russian will minimally require several years of heavy experience with Russian. Rather, what might be required is years of exposure to *easy* L2 input (the definition of easy input, of course, changing over time). Much of the exposure to Russian experienced by the woman in the case study was situated in daily home life. Such speech—often clearly tied to the concrete, present situation; often rooted in shared experiences; and typically benefiting from extensive shared knowledge—may, in general, be easier to process than the speech experienced by many of the experimental participants (e.g., university lectures, language used in service encounters, or language used in casual social situations with adult Russian acquaintances.)

4.7.5. Pienemann's Processability Theory

Another model which may have points of contact with the approach taken in this dissertation is Pienemann's (1998) Processability Theory. He basically adopts Levelt's (1989) model of speech production, which was in turn based on Kempen and Hoenkamp's Incremental Production Grammar, which was in turn based on Bresnan's theory of Lexical Functional Grammar (LFG) (see Pienemann, 1998, for references).

LFG combines a constituent-structure component with a component which specifies feature-structures (F-structures) of sentences. Constituent structures display the hypothesized organizational interrelationships within and among the phrasal components of a sentence. Feature-structures are intended to reflect something closer to the semantic interrelationships of the components of a sentence, making use of lexical content and abstract grammatical features (such as subjecthood, objecthood, lexical categories, and grammatical features such as person and number). In constructing a feature structure, features originating in various parts of a sentence must unify in the ultimate F-structure. The concept of feature unification might be considered similar in spirit to the concept of cue convergence.

Pienemann attempts to use feature unification as a basis for predicting processing difficulty, and in turn, predicting acquisition difficulty. He reasons that it should be easier to pass features between words than between phrases, and easier between phrases than between clauses. Pienemann attempts to use this hierarchic pattern of feature passing to account for acquisition of word order in L2 German (familiar from the earlier Multidimensional Model—Meisel, Pienemann and Clahsen, 1981). He also applies the hierarchy to the acquisition of grammatical morphology: word-related morphology (e.g., grammatical number) ought to be learned at an early stage, while morphology that requires that features unify within phrases (e.g., noun-adjective agreement) should be acquired later, and morphology which requires that features unify across separate phrases (e.g. subject-verb agreement) should be acquired still later.

There is some question regarding whether Pienemann's model is dealing with processing in the psycholinguistic sense. There are times when the internal grammar that Pienemann seems to have in mind is partly distinct from the processing mechanisms, which according to him, can differ widely depending on the modality of language processing (e.g., reading comprehension vs. listening comprehension).

As to the specific predictions regarding the acquisition order of inflectional morphology, they do not seem consistent with the pattern of development of sensitivity to inflection in L2 Russian as observed in the previous chapter. In the striking case of verb inflection in Experiment 2, the sentence level process of subject-verb person agreement appears to be the most readily acquired feature of verbal inflection, if not of L2 Russian inflectional morphosyntax in general. By contrast, sensitivity to aspect inflection, arguably an example of word-level morphology in Pienemann's approach, is considerably less readily acquired.

One component of the research reported by Pienemann would appear to be related to the experiments of Chapter 3. Specifically, Pienemann reports an experiment involving sentences with inflectional errors (errors of subject-verb person agreement). The experimental task was sentence matching (SM), rather than error detection. In SM, participants are required to decide whether two visually presented sentences are the same or different. Native speakers are known to respond more slowly in the case of ungrammatical sentences than in the case of grammatical sentences. Pienemann found that early nonnative speakers show little sensitivity to the ungrammaticality involved in person agreement, while skilled L2 users show a grammaticality effect similar to that seen with native speakers. There is no indication of how long the skilled L2 learners had been learning German or in what contexts

4.7.6. Information Processing Models.

A popular genre of SLA models sometimes goes under labels such as *cognitive approaches* or *information processing approaches*. To some extent the popularity of these approaches may derive from the important role they appear to assign to language pedagogy, including long-standing practices of "rule" teaching followed by controlled production practice. Within this genre of SLA models one finds work by McLaughlin (1987, Chapter 6, 1978, 1990; McLauglin, Rossman & McLeod, 1983; McLaughlin & Roberto, 1996), Bialystok (1978; 1990), Schmidt (1990, 1992; Schmidt & Frota, 1986), Hulstijn (1990), Towel & Hawkins (1994) and Skehan (1998). Such models may in part be inspired by a subjective sense that early language use is extremely effortful, while fluent nonnative speakers seem to function more similarly to native speakers in that for them, speaking does not seem to require a lot of concentration. Given this apparent change from effortful language use to language use that is automatic (at least in informal terms—see footnote 24), it could appear that research findings related to the development of automaticity might provide insights into SLA.

A possible weakness of this overall group of approaches is a tendency to under-emphasize what exactly it might be that is being automatized. That is, scholars pursuing these approaches draw heavily on certain areas of cognitive psychology, while devoting less attention to research related to language processing. For example, although Skehan (1998) has a chapter entitled "Psycholinguistic Processes in Language Use and Language Learning," the chapter contains mainly references to SLA research or psychological research in the information-processing tradition, rather than the psycholinguistic tradition. The brief discussion of comprehension processes in Chapter 1 draws, in a limited, way on Clark & Clark (1977).

This lack of attention to the nature of the processes allegedly becoming automatized is apparent from McLauglin's early applications of the concept of controlled and automatic processing to SLA (e.g., McLaughlin, 1978). He drew on the work on controlled and automatic information processing of Schneider and Shiffrin (1977; Shiffrin & Schneider, 1977). That research involved a number of experiments in which participants pressed a key in response to seeing a letter or digit belonging to a set they

had trained on. The general thrust of the findings was that a lot of thought went into early decisions regarding whether a particular character was in the training set, whereas after many trials, no apparent thought went into reactions, and in fact, detections and responses to characters that were "automatized" in one experimental task were difficult to repress in a subsequent experimental task.

In order to apply these findings to language learning, it would be minimally necessary to determine which aspect of language processing corresponds to the visually displayed characters in the experiments, and which aspect of language processing corresponds to the act of pressing the key in the experiments. In McLaughlin (1987), it is clear that the model of learning (i.e., changing from controlled to automatic processing, possibly including restructuring of the processes along the way) is intended as a possible account of comprehension learning as well as production learning. But what in comprehension corresponds to the training set of characters in Schneider and Shriffrin's experiments, what corresponds to the visual detection of members of the training set, and what corresponds to the response, that is the pressing of the key? Without some idea of where the analogues lie, the claim that learning involves automatization is not too enlightening. Automatization of what?

McLaughlin in fact placed more emphasis on the role of information-processing mechanisms in spoken production than in comprehension. He cites an early model of Levelt's (1978) which appears to decompose speech production into a set of sub-modules. The system starts with "intentions", relates them to syntactic schemata, constructs phrases and so on, in what appears to be a strictly serial manner. Here again, the details regarding the manner in which specific aspects of linguistic form become automated in production is left largely to the imagination. We can imagine that the application of metalinguistic knowledge somehow constitutes the "attentional control" of the early stages of learning. In the course of such production, the attentional control manages to coactivate nodes (in McLaughlin's terminology) that will later be activated directly by the input. This input presumably consists in various aspects of the conceptual flow that is being verbalized. The nodes activated are presumably those that trigger spoken production. Note that once the inputs are directly associated with the production nodes, what has formed has no relationship to the metalinguistic knowledge involved in the attentional control that coactivated inputs and nodes in the first place. For example, a pedagogical rule (such as a rule under whose guidance the learner chooses the accusative form of a noun to mark a direct objects) is not what becomes automated. Rather, the strategy of applying that pedagogical rule fosters the activation of the right nodes in the presence of the right input. In the automatic stage, the input-to-nodes relationship has taken on a life of its own, and the pedagogical rule may or may not have continued to exist in long term memory as a piece of metalinguistic knowledge.

From the standpoint of this dissertation, the whole idea of developing linguistic production mechanisms via the controlled application of pedagogical rules is problematic. It certainly depends on a theoretical understanding of grammar as information regarding what goes where in sentences, rather than as a system of processing cues which are not linguistically acquired unless their functions are acquired. If the functions of grammatical elements are fundamentally comprehension functions, then is it realistic that learners could automatize their production of forms without acquiring the functions of those forms? Again, it would appear that some other type of non-linguistic knowledge is a more reasonable candidate for explaning increases in the nativelikeness of spoken output in the absence of increases in the nativelikeness of the corresponding comprehension mechanism. Truscott (1998) reasons that

> ...[metalinguistic] knowledge could, in certain circumstances, serve as a supplement to competence. As it becomes automatized, speakers might come to use it fluently, possibly making up for weaknesses in competence (p. 125)

Nevertheless, it is not obvious that such mechanisms could assume the major role in L2 comprehension or production. Perhaps a subsidiary role is more plausible.

4.7.7. Schmidt's Noticing Hypothesis

Another highly influential concept often associated with information-processing approaches (and other so-called cognitive approaches) is Schmidt's Noticing Hypothesis (Schmidt & Frota, 1986; Schmidt, 1990). This hypothesis assigns a role to attention, if not attentional control, in the early stages of second language learning. Schmidt's experiences learning Portuguese indicated that a learner can know about a form and be exposed to it many times without being aware of its presence in the input. The original form of the hypothesis held that the learner would not start to use such a form in his or her spoken production until s/he started to notice it during comprehension. From the perspective of the research reported in Chapter 3, it would appear that at least in the case of much Russian inflection, simply hearing an inflectional form, such as *br'ukax*, 'pants:locative:plural' was not the same as hearing it and reacting to it as though it were something other than a neutral alternative form of the word *br'uki*, 'pants:nominative:plural'. That is, "noticing" would have to involve both detection of the form and relevant reaction to the form as a processing trigger. The frequent detection of the form and relevant reaction to it as a processing trigger would lead to the association of the form with the corresponding contextual conditions.

Schmidt's noticing stage would fit nicely at the point where there is a reaction to some inflectional form, basically, phase B in the hypothesized six phases discussed early in this chapter. We might imagine that in the case of a linguist-language learner such as Schmidt phase B "surprise" experiences would lead to frequent analytical reflections. However, the phenomenon of grammatical form "coming to the notice" of the processor, that is creating a need for processing in the absence of the ability to perform the processing would be more general.

4.8. THE LONG ROAD AND RESEARCH DATA

I have argued that L2 Russian inflectional learning involves developments which require a few years of learning in most cases, and that metalinguistic strategies can contribute to the impression of nativelike production in the shorter range. If this is correct, then researchers investigating some aspect or other of L2 development in a particular target language might well keep two questions in mind: 1) Can the mechanisms involved in processing the grammatical element of interest develop in learners (in a truly linguistic sense) within the time-frame involved in the proposed study? 2) Does the instrument being used to detect development detect on-line linguistic processes, the application of metalinguistic strategies, or some mix of the two?

4.8.1. The time-frame of research projects

Schachter (1998) raises the issue of the overall time-for-testing involved in connection with computer-controlled SLA experiments:

> Computer-controlled language learning studies vary widely in amount of training, from an hour or less to 3 to 5 weeks. Yet the results are often discussed with no mention of this major difference or how it might affect their interpretation. (p. 567)

She goes on to claim that

> ...with language structures fairly difficult to learn (pseudo cleft, adverb fronting with subject/aux inversion, dative movement and some derivational morphology in French, relative clauses, complex wh-questions) shorter training periods favor the explicitly trained groups, midrange training provides mixed results and longer training periods allow the implicitly trained groups time to catch up. (p. 570)

Schachter expresses special skepticism regarding hour-or-less learning experiments. However, in the case of Russian inflectional morphology, even if effects were to be measured that resulted from several weeks of treatment, there might be grounds for questioning what variety of learning was being detected (see 4.6 above).

On the other hand, as noted, considerable learning of *some* components of L2 Russian does apparently occur relatively early (for example, accounting for the high performance seen in the concurrent picture-selection sub-task of the Meaning-Oriented task). Thus, in addition to being relevant more to certain types of learning (as discussed in section 4.6) than to others, research projects involving short-term learning may also be more relevant to certain aspects of the L2, than to others.

In order to decide whether some feature of a language is amenable to investigation within a short time-frame, it could be useful to independently establish the time-frame of acquisition of that feature, for example, by determining how early learners show on-line sensitivity to manipulations affecting it. Given the findings of the previous chapter, I would argue, no one could meaningfully study the effects of a particular pedagogical treatment on most aspects of L2 Russian inflectional learning over the course of a few weeks, or even a semester or two.

This consideration may have implications for learning projects involving artificial miniature languages. For instance, Yang & Givón (1997) proposed quite plausibly that grammar would be acquired more readily in the presence of familiar vocabulary than in the presence of unfamiliar vocabulary. The artificial language employed (Kecki) included a range of inflectional categories marked by suffixes. Among these were subject agreement and tense-aspect inflection. Post-treatment measures of grammar learning included a recall test, grammaticality judgements to written sentences, picture descriptions, and a translation task. Although in the report, the results are not broken down by grammatical phenomena, the various measures did detect clear learning. However, it is one thing to detect learning. The more important concern should be whether the learning that is measured involves the development of language processing mechanisms. We have grounds to be skeptical, even if we take Russian to be an extreme case. How much opportunity would there be in fifty hours for tense-aspect processing to develop? Experiments such as those reported in Chapter 3, but based around a language with more nearly one-to-one form-function relationships in the inflectional systems, might give us an indication of the reasonableness of a fifty-hour learning experience as the basis for the acquisition of aspectual processing.

Classroom-based studies may be subject to the same problem. For example, White (1991a,b) looked for changes in the L2 English of a group of learners in the matter of their tendency to place adverbs between transitive verbs and direct objects (e.g., *Mary*

watches often television.). She compared learners who received only incidental exposure to positive evidence with learners who were provided with explicit negative evidence. The treatment took place over a two week period. Recall that in Chapter 1, I argued that what is unacquired by learners who produce utterances such as *Mary watches often television* is a parsing strategy (which takes the post-verbal adverb as an extremely reliable cue that there will be no direct object). It would not be surprising to find that some parsing strategies take years to develop. It is interesting, though, that in relation to the same aspect of L2 English, Trahey and White (1993) did manage to effect a change in the spoken production of learners in the course of two weeks through an "input flood" of speech containing the target word order pattern. This result was somewhat discounted because it was not categorical. That is, the learners produced a mix of target-like and non-target-like utterances. Given that Trahey and White were looking for a categorial change, they took these results negatively. However, in processing terms, it is feasible that the input flood had in fact caused the *onset* of the relevant learning. Over an extended period, the comprehension system might become increasingly nativelike, thus increasingly training the production system to be nativelike. There could be a problem however, in that it might take a continuation of the input flood to sustain the rate of development which was initiated.

Thus, Trahey and White's (1993) findings might illustrate not only the potential limitations, but also the potential relevance, of short-term classroom studies which compare pedagogical treatments. On the one hand, such studies may in principle be unable to demonstrate ultimate learning of those grammatical elements which requires extensive, long-term experience with a language. On the other hand, given a viable theoretical account of those long term developments, it might be possible to demonstrate that a particular classroom approach creates conditions which should facilitate long-term learning. It was argued in Chapter 3 that the conditions that distinguished the Meaning-Oriented task from the Form-Oriented task in fact increased the tendency for inflectional forms to be processed. These differences included increased familiarity with the texts and longer pauses after text segments. Assuming that learning is ultimately a function of the amount of processing of the relevant variety, then such a result could be said to be of pedagogical relevance. However, the scale of experience on which the effects are realized might be such as to eliminate the possibility of directly testing the benefits of these properties of the input on the ultimately desired developments. Rather we could only extrapolate from an understanding of long-term developments to decisions regarding short-term treatments. That might be a disappointing result for pedagogues who desire proof that a particular treatment will produce the desired long-term effect. However, it may well be the best that nature has to offer. From the standpoint of full-blown SLA (see the case-study earlier in this chapter), even what is considered long-term by pedagogical standards may be relatively short-term in relation to particular developments.[34]

In general, it might be wise to exercise caution in attributing *any* learning to normal language mechanisms if that learning occurs in a particularly brief timeframe. Even if learning resulting from focused input processing (e.g., VanPatten & Cadierno, 1993) is shown to be more effective than learning resulting from rule teaching plus production practice, that may primarily indicate that the former approach is superior to the latter in terms of fostering metalinguistic knowledge that can be used in planned production. It may also be that the input processing leads directly to learning of a more purely linguistic variety. That would need to be demonstrated by examining whether the

[34] On the other hand, some language programs may, for socio-cultural reasons, place a high value on the ability of learners to simulate nativelike production through metalinguistic strategies.

target items or structures have begun to function on-line in comprehension or remain inert.

4.8.2. Varieties of data

Finally, the results reported in Chapter 3 could have some bearing on the relevance of different varieties of data employed in second language research. If the radical difference between the Meaning-Focused task and the Printed-Sentences task reflects a difference in the nature of on-line processing and off-line metalinguistic analysis, then perhaps these variables should be routinely separated in research employing grammaticality judgements. Davies and Kaplan (1998) found evidence that in making grammaticality judgments under unconstrained conditions, language learners often appeal to the sort of metalinguistic reasoning that is specific to the instructed L2 environment. The closest equivalent in strategies employed native speakers in making grammaticality judgments in Davies and Kaplan's study was the occasional (attempted) application of prescriptive grammatical knowledge. However, it played a relatively small role in L1 grammaticality judgments in comparison with the role played by metalinguistic analysis in the L2 grammaticality judgments. Similarly, the native controls in the Meaning-Oriented task of Chapter 3 appeared to be reacting to processing clashes, not to metalinguistic knowledge about case usage or aspect usage. In fact, the linguist's notion of "ungrammaticality" might arguably be redundant in the context of an adequate understanding of processing. Excluding judgments related to prescriptive grammatical training, it seems reasonable to expect that native-speaker ungrammaticality reactions are reactions to incompatible processing cues (including parsing failures). Thus for native-speakers, it may be of little consequence whether judgements are made to written sentences without time constraints, or to spoken utterances under time constraints. For nonnative speakers these two different conditions may be detecting relatively distinct underlying realities: normal comprehension processes vs. (normal comprehension processes plus) the application of metalinguistic knowledge in an analytical, problem-solving mode.[35]

For similar reasons, we have noted that the use of spoken production data from instructed learners may provide a mixed picture, arising from relatively distinct underlying processes. That is, when nonnative speakers are in a position to plan ahead, employing metalinguistic knowledge, the results may be different from what would be attributable to purely linguistic speech production processes. This might be true to a smaller extent with non-instructed learners, given the tendency of learners from puberty on (if not earlier) to make metalinguistic observations (Felix, 1987).

---

[35] White (1981) administered her stimuli for grammaticality judgments under time pressure, providing only a four-second interstimulus interval. Her purpose was to prevent reliance on translation into the L1 as the basis for the grammaticality judgments. She was not specifically concerned to ensure that reactions to ungrammaticality were based on on-line reactions rather than off-line reflections. However, I would argue that any time-constrained grammaticality judgment task with auditory stimuli would also have the latter effect. Furthermore, given the apparent limited ability of L2 listeners to maintain grammatical details of entire sentences in short-term memory for methalinguistic analysis, it may be that even without time constraings, an L2 grammaticality judgment task with auditory stimuli will tend to reflect effects of on-line processing, while a parallel task with written stimuli will additionally reflect effects of other varieties of learning.

4.9. WRAP-UP

This dissertation was in part inspired by a desire to better understand the recalcitrance ("fossilizability") of some aspects of second language systems, a topic that is commonly discussed in relation to spoken production, and not in relation to comprehension. Selinker's (1972) foundational article dealt with the frequent lack of permanent pedagogical success in eliminating certain patterns from the spoken production of learners. Similarly, White (1996) talked about the role of negative evidence in terms of whether it could "cause the elimination of the ungrammatical word order" (White, 1996, p. 101), as evidenced in spoken production.

What occurred to me was that perhaps we are looking for the causes of arrested development in the wrong place. Perhaps development is getting stuck, not in the system responsible for the form of spoken output, but rather in the listening comprehension system. For example, perhaps when learners of L2 English—despite repeated teaching, drilling and overt corrections—continue to produce English sentences of the form *My brother work at the Seven-Eleven*, we might find, if we were to look, that verbal inflectional form is still inert within the comprehension system of such learners. Since most models of SLA place substantial emphasis on the role of input, we should perhaps look more closely at the extent to which various aspects of input are in fact doing anything at all inside the learner. In short, if the functions of many commonly unacquired features of L2s are primarily comprehension functions, then we are unlikely to understand the nonacquisition of those features from looking at production data alone, nor from the results of off-line grammaticality judgment tasks.

Russian was chosen as the language for investigation largely on the basis of widespread reports of long-term inflection-related struggles experienced by adult learners. It was hoped that looking at a language that promised numerous prolonged developments would provide increased opportunities for detecting patterns in the data. The challenge was to determine whether various inflectional forms were active or inert within the listening comprehension system of learners. The experimental technique was chosen with a view to practicality (and with a sense of the need to gather data from a substantial number of participants and that was related to a moderately wide range of phenomena). I feel that in the process I have demonstrated that in principle, the investigation of L2 comprehension processes is not inordinately difficult. The fact that L2 morphosyntax has not more frequently been investigated from the vantage point of comprehension processes may reflect a cultural bias which understands language learning as "learning to speak" and views target language grammars as sets of recipes for constructing sentences.

Research questions are always asked within a framework of assumptions. The framework of assumptions I chose was drawn from psycholinguistic research on language comprehension. In Chapter 1, we saw that the range of processes involved in normal language comprehension might adequately handle the needs of L2 listeners at all stages, but with the relative reliance on various components shifting over time. For example, in early learning, lexical meanings (including selectional requirements) and inferential mechanisms might play a proportionately larger role, while grammatical cues might play a proportionately smaller role, than would be the case in later L2 use. Thus, with comprehension, there is no obvious role for cognitive processes that fall outside the pale of normal language comprehension processes. With language production, this is not so clear. In fact, there appears to be some scope for the application of metalinguistic, analysis-based planning in L2 production, while such processes typically have little or no role in normal L1 production. This difference between L2 comprehension and production may point us in the direction of a plausible account of the reemergence of nonnative

features that were seemingly eradicated. What may reemerge are production patterns that do not in any way clash with the learners' current comprehension mechanisms, while features of the target language that as yet have no function for the learners—or that function only variably and weakly—vanish. The output improvements observed earlier would have been due to nonlinguistic production strategies. Those strategies might cease to be employed for various reasons. This would give the appearance of regression in learning. The comprehension system, however, would not have gone through the same pattern of apparent progression and regression.

Happily, the L2 comprehension system appears to develop rather quickly in some areas. Adequate sound discrimination and word detection mechanisms must develop quickly and certain aspects of lexical functioning, such as the exploitation of argument structure or selectional restrictions, may come into play fairly readily, as may certain aspects of word order (e.g., postpositional order as opposed to prepositional order in the case of adpositions in some languages, for example). Within a few hundred hours of exposure, a considerable amount of comprehension can already be underway. Recall from Chapter 1 the image of a "wall of sound" that I used to describe an encounter with a language totally unknown to me. When a language is heard for the first time it is heard as "continuous noise and not as words" (Carroll, 1999, p. 357). Recall also from Chapter 1 the enormity of the native processes that convert the wall of sound into an unnoticed window to meaning. Recall also the speed at which those processes operate. When nonnative listeners, a few months into their language learning, are able to comprehend considerable stretches of speech at normal speeds, it should be clear that some significant developments have already occurred. It may be that, at first, just enough new learning takes place to allow the L2 comprehension system to tap into L1 mechanisms. In any case, it is clear that *something* develops relatively quickly. A self-aware, psycholinguistically literate second language learner is justified in marveling at how soon the wall of sound starts to serve as a window—if perhaps a smoked-up window—to meaning. But while some aspects of L2 processing ability appear to develop fairly readily, other aspects of L2 processing ability appear to require thousands of hours of experience with a live language in order for them to develop. The difference in ease of acquisition could be rather straightforwardly related to the time-stability of the conceptual component or conceptual process with which linguistic forms must become associated in order for learning to be successful.

Sadly, the processes that appear to require longer-term developments may be less amenable to our direct modification than we might wish. On a personal level, I can recall the considerable frustration (to say nothing of humiliation) that I experienced over the tendency of my L2 Russian production system to ignore the sex of referents when supplying third person singular pronouns. (Russian, like English, has masculine and feminine forms, in the nominative *on* [on], 'he' and *ona* [ʌ'nɑ], 'she'.) It was interesting to observe introspectively that my comprehension system also did not make the relevant distinction. That is, if I began listening in the middle of a text in which a woman was being referred to by pronouns, I might readily start constructing a mental model with a male referent. In one case that I documented in my language learning diary, a narrative participant was referred to as a *kontrol'or*, 'inspector', a noun which could refer to either a man or a woman. In spite of the fact that the person was subsequently referred to with feminine pronouns and feminine verb agreement, I did not replace the male inspector-token in my mental model with a female inspector-token until after several listenings to the same text. When during the second and third years of my Russian learning I realized that this problem was still serious (despite my longstanding metalinguistic understanding of the relevant issue), I designed pedagogical self-treatments. One of my exercises gave

me literally thousands of opportunities to attach the appropriate pronouns to appropriate real-world referents as they moved along the opposing escalator going into or coming out of the St. Petersburg Metro. Yet I continued to have it pointed out to me regularly that I had, for example, referred to my wife as *on*, 'he'. Recently, well over four years into my Russian learning, I was listening to a children's legend in which a snake entered the story at a certain point and interacted for a while with the main protagonist. Well into the segment of the story in which the snake was participating, I realized that I was envisioning a "girl snake". I conjecture that the reason was that the snake was being referred to by feminine pronouns and verb agreement, though I had no overt awareness of the fact. Perhaps my comprehension system is finally now attending to the gender difference.[36] Hopefully, the humiliation of being told to refer to my wife as *she* rather than *he* will "soon" be relegated to the dustbins of my language learning history.

A similar anecdote involves the Finnish woman discussed in the case study earlier in this chapter. She described a stage in her own Russian learning when her speech production system supplied feminine agreement for second person referents, regardless of their sex, and masculine agreement for first person singular, in spite of the fact that she knew that she was female. She possessed the relevant metalinguistic knowledge to understand that the mistake was a mistake, but her language processing system, apparently having a mind of its own, seemed to be using gender morphology as a marker of person. She speculated that the reason for this was that her overwhelming interactional experience was with her Russian husband, who of course referred to her (second person) with feminine forms and to himself (first person) with masculine forms. As gender and person were thus heavily confounded in her input, her comprehension system latched onto gender marking as a reliable cue to person, and her production system started to follow suit, against her wishes and against her solid metalinguistic understanding. The problem eventually passed.

In the end, do learners of L2 Russian acquire much or most of the inflectional system for use in comprehension? Perhaps most learners will not, due to the fact that the process of acquisition requires a larger amount of experience with Russian as a spoken language than they will ever have. For those who adopt Russian as their primary language of life for a number of years, the chances of full (or nearly full) development might be greater. To demonstrate that this is so, we would need to gather data from a reasonable sample of such learners.

The prolonged nature of the acquisition of inflection in L2 Russian may be due in part to the complexity of Russian inflection, and in particular, to the existence of multiple inflectional classes and the many-to-many nature of the form-to-function relationships. To demonstrate that this is a cause of the protracted learning pattern, we would need to compare the acquisition of grammatical features such as case and aspect in L2 Russian with the acquisition of analogous inflectional categories in other languages where the form-to-function relationships are more nearly one-to-one. In any such investigations, it should be recognized that without looking at what is happening during online comprehension of speech, we cannot construct a revealing picture of the development of the L2 system.

---

[36] Unfortunately, a snake is referred to by feminine pronouns and verb agreement, regardless of whether it is male or female!

References

Alpatov, V.M. Chto sleduet schitat' isxodnoj edinitsej morfologicheskogo opisanija? In Bogdanova, S.I. & Gerda, A.S. (Eds.) *Morfemika: Printsipy Segmentatsii, Otozhdestvlenija i Klassifikatsii Morfologicheskix Edinits*, pp. 4-11. St. Petersburg: St. Petersburg University.

Altmann, Gerry T.M. (1997). *The Ascent of Babel: An Exploration of Language, Mind and Understanding*. Cambridge: Cambridge University Press.

Altmann, Gerry T.M. (1999). Thematic role assignment in context. *Journal of Memory and Language, 4*, 124-45.

Andersen, Roger W. (1991). Developmental sequences: The emergence of aspect marking in second language acquisition. In Huebner, Thom & Ferguson, Charles A. (eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theories*, pp. 305-24. Amsterdam: Benjamins.

Andersen, Roger W. & Shirai, Yasuhiro. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition 16*, 133-56.

Andersen, Roger W. & Shirai, Yasuhiro. (1996). The primacy of aspect in first and second language acquisition: The pidgin-creole connection. In Ritchie, William C. & Bhatia, Tej K. (eds.), *Handbook of Second Language Acquisition*, pp. 527-70. San Diego: Academic Press.

Anderson, John R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406.

Anderson, John R. (1983). *The Architecture of Cognition*. Cambridge, Mass.: Harvard University Press.

Anderson, Stephen R. (1985). Inflectional morphology. In Shopen, Timothy (ed.) *Language Typology and Syntactic Descriptions: Volume III, Grammatical Categories and the Lexicon*, pp. 150-201. Cambridge: Cambridge University Press.

Anderson, Stephen R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University press.

Antinucci, Francisco & Miller, Ruth. (1976). How children talk about what happened. *Journal of Child Language, 3*, 167-89.

Archibald, John & Libben, Gary. (1995). *Research Perspectives on Second Language Acquisition*. Toronto: Copp Clark.

Aronoff, Mark. (1994). *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, Mass.: MIT Press.

Bailey, Nathalie, Madden, Carolyn, & Krashen, Stephen D. (1974). Is there a "natural sequence" in adult second language learning? *Language Learning, 24*, 235-43.

Bardovi-Harlig, Kathleen. (1992). The relationship of form and meaning: A cross-sectional study of tense and aspect in the interlanguage of learners of English as a second language. *Applied Psycholinguistics 13*, 253-78.

Bardovi-Harlig, Kathleen. (1994). Anecdote or evidence? Evaluating support for hypotheses concerning the development of tense and aspect. In Tarone, Elaine E., Gass, Susan M., & Cohen, Andrew D. (eds.), *Research Methodology in Second-Language Acquisition*, pp. 41-60. Hillsdale, N.J.: Erlbaum.

Bardovi-Harlig, Kathleen. (1998). Narrative structure and lexical aspect: Conspiring factors in second language acquisition of tense-aspect morphology. *Studies in Second Language Acquisition 20*, 471-508.

Bardovi-Harlig, Kathleen. (1999). From morpheme studies to temporal semantics: Tense-aspect research in SLA. *Studies in Second Language Acquisition 21*, 383-420.

Bates, Elizabeth & MacWhinney, Brian (1981). Second language acquisition from a functionalist perspective: Pragmatic, semantic and perceptual strategies. In Winitz, Harris (ed.), *Annals of the New York Academy of Sciences Conference on Native and Foreign Language Acquisition*, pp. 190-214. New York: New York Academy of Sciences.

Bates, Elizabeth & MacWhinney, Brian. (1982). Functionalist approaches to grammar. In Wanner, Eric & Gleitman, Lila (eds.), *Language Acquisition: The State of the Art*, pp. 173-218. Cambridge: Cambridge University Press.

Bates, Elizabeth & MacWhinney, Brian. (1989). Functionalism and the Competition Model. In MacWhinney, Brian & Bates, Elizabeth (eds.), *The Cross-Linguistic Study of Sentence Processing*, pp. 3-73. Cambridge: Cambridge University Press.

Beard, Robert. (1995). *Lexeme-Morpheme Base Morphology*. Buffalo: SUNY press.

Berko, Jean. (1958). The child's learning of English morphology. *Word 14*, 150-77.

Berman, Ruth A. & Slobin, Dan Isaac. (1994). Becoming a proficient speaker. In Berman, R.A. & Slobin, D.I. (eds.), *Relating Events in Narrative: A Cross-Linguistic, Developmental Study*, pp. 597-641. Hillsdale, N.J.: Erlbaum.

Bialystok, Ellen. (1978). A theoretical model of second language learning. *Language Learning, 28*, 69-84.

Bialystok, Ellen. (1990). *Communication Strategies: A Psychological Analysis of Second Language Use*. Oxford: Blackwell.

Bialystok, Ellen. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition, 16*, 157-68.

Bialystok, Ellen & Ryan, Ellen Bouchard. (1985). Toward a definition of metalinguistic skill. *Merrill-Palmer Quarterly, 31,* 229-51.

Birdsong, David. (1989) *Metalinguistic Performance and Interlinguistic Competence.* Berlin: Springer-Verlag.

Bley-Vroman, Robert. (1989). What is the logical problem of foreign language learning? In Gass, Susan M. & Schachter, Jacquelyn (eds.), *Linguistic Perspectives on Second Language Acquisition,* pp. 41-68.

Blekher, Marina. (2000). *Word-Type Effects in the Lexical Processing of Russian-English and French-English Bilinguals.* University of Alberta Doctoral Dissertation.

Bloom, Lois, Lifter, Karin, & Hafitz, Jeremie. (1980). Semantics of verbs and the development of verb inflection in child language. *Language, 56,* 386-412.

Bloomfield, Leonard. (1933). *Language.* New York: Holt.

Bock, Kathryn & Levelt, Willem. Language production: Grammatical encoding. In Gernsbacher, Ann (ed.), *Handbook of Psycholinguistics.* San Diego: Academic Press.

Bolinger, Dwight. (1975). Meaning and memory. *Forum Linguisticum 1,* 2-14.

Bondarko, A.V. (1997). O sisteme analisa semantiki glagolnogo vida (na materiala russkogo jazyka. In Chertkova, M. Ju. (ed.), *Trudy Aspektologicheskogo Seminara Filologicheskogo Fakul'teta MGU Im. M.V. Lomonosova 1, Tom 3,* pp. 6-18. Moscow: Moscow University Press.

Bondarko, L.V. (1998). *Fonetika Sovremennova Russkovo Jazyka.* St. Petersburg: Petersburg University Press.

Boots-Ebenfield, Marc. (1995). *The Pragmatics of Grammatical Aspect in the Interlanguage of American College Students of Russian: An Analysis of Oral Proficiency Interviews Before and After Study Abroad.* Ph.D. dissertation, Bryn Mawr College.

Bransford, John D., Barclay, J. Richard, & Franks, Jeffery J. (1972). Semantic memory: A constructive versus interpretive approach. *Cognitive Psychology, 3,* 193-209.

Brecht, Richard D. (1985). The form and function of aspect in Russian. In Flier, Michael S. & Brecht, R. D. (eds.), *Issues in Russian Morphosyntax.* Columbus: Slavica.

Brown, Cynthia. (2000). The interrelation of speech perception and phonological acquisition from infant to adult. In Archibald, John (ed.), *Second Language Acquisition and Linguistic Theory.* Oxford: Blackwell.

Brown, Nicholas J. (1996). *Russian Learner's Dictionary: 10,000 Words in Frequency Order*. London: Routledge.

Brown, Roger. (1973). *A First Language: The Early Stages*. Cambridge, Mass.: Harvard University Press.

Brown, Roger & Fraser, Colin. (1963). The acquisition of syntax. In Cofer, Charles N. & Musgrave, Barbara S. (eds.), *Verbal Behavior and Learning: Problems and Processes*, pp. 158-201. New York: McGraw Hill.

Brown, Roger, Fraser, Colin & Bellugi, Ursula. (1964). Explorations in grammar evaluation. In Belugi, Ursula & Brown, Roger (eds.), *The Acquisition of Language*, pp. 79-92. *Monographs of the Society for Research in Child Development*, 29.

Bybee, Joan L. (1985). *Morphology: A Study in the Relation Between Meaning and Form*. Amsterdam: Benjamins.

Bybee, Joan L. (1991). Natural morphology: The organization of paradigms and language acquisition. In Huebner, Thom & Ferguson, Charles A. (eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theories*, pp. 67-91. Amsterdam: Benjamins.

Bybee, Joan, L., Pagliuca, William, & Perkins, Revere. (1994). *The Evolution of Grammar: Tense Aspect and Modality in the Languages of the World.* Chicago: University of Chicago Press.

Bygate, Martin. (1996). Effects of task repetition: Appraising the developing language of Learners. In Willis, Jane & Willis, Dave (eds.), *Challenge and Change in Language Teaching*, pp. 136-46. London: Heinemann.

Carlson, Greg N. Tannenhaus, Michael K., & Trueswell, John C. The role of thematic structure in interpretation and parsing. *Language and Cognitive Processes, 4*, 211-34.

Carroll, John & Sapon, S.M. *Modern Language Aptitude Test*. New York: Psychological Association.

Carroll, Susanne E. (1999). Putting 'input' in its proper place. *Second Language Research, 15*, 337-88.

Cazden, Courtney. (1968). The acquisition of noun and verbs inflections. *Child Development*, 39, 433-48.

Chaudron, Craig. (1983). Research on metalinguistic judgments: A review of theory, methods, and results. *Language Learning, 33*, 343-77.

Chen, Hsuang-Chih & Leung, Yuen-Sum. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 316-25.

Chomsky, Noam A. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT press.

Chomsky, Noam A. (1972). Some empirical issues in the theory of transformational grammar. In Stanley Peters (ed.), *Goals of Linguistic Theory*, pp. 63-130. Englewood Cliffs: Prentice Hall.

Chomsky, Noam A. (1981). *Lectures on Government and Binding*. Dordrecht: Reidel.

Chomsky, Noam A. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, Nom A. (1986). *Knowledge of Language*. New York: Praeger.

Chvany, Catherine V. (1985). Backgrounded perfectives and plot line imperfectives: Toward a theory of grounding in text. In Flier, Michael S. & Timberlake, Alan (eds.), *The Scope of Slavic Aspect*, pp. 247-73. Columbus: Slavica.

Clark, Eve V. (1998). Morphology in language acquisition. In Spencer, Andrew & Zwicky, Arnold M. (eds.), *The Handbook of Morphology*, pp. 375-89. London: Blackwell.

Clark, Herbert H. & Clark, Eve. (1977). *Psychology and Language*. New York: Harcourt, Brace & Jovanovich.

Comrie, Bernard. (1976). *Aspect*. Cambridge: Cambridge University Press.

Comrie, Bernard. (1985). *Tense*. Cambridge: Cambridge University Press.

Comrie, Bernard. (1991). *Language Universals and Linguistic Typology, 2nd ed.* Chicago: University of Chicago Press.

Cook, Vivian James. (1993). *Linguistics and Second Language Acquisition*. New York: St. Martin's Press.

Coorman, Ann & Kilborn, Kerry. (1991). Functionalist linguistics: Discourse structure and language processing in second language acquisition. In Huebner, Thom & Ferguson, Charles A. (eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theories*, pp. 195-224. Amsterdam: Benjamins.

Coppieters, Rene. (1988). Competence differences between native and non-native speakers. *Language, 63*, 544-73.

Corder, S. Pit. (1967). The significance of learner errors. *International Review of Applied Linguistics, 5:*161-9.

Croft, William. (1990). *Typology and Universals*. Cambridge: Cambridge University Press.

Crookes, Graham. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition, 11*, 367-83.

Cutler, Anne, Mehler, J., Norris, Dennis, & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology, 24*: 381-410.

Dahl, Östen C. (1985). *Tense and Aspect Systems*. Oxford: Blackwell.

Davies, William D. & Kaplan, Tamar, I. (1998). Native Speaker vs. L2 Grammaticality Judgements. *Applied Linguistics, 19*, 183-203.

de Villiers, Jill, & de Villiers, P. (1973). A cross-sectional study of the development of grammatical morphemes in child speech. *Journal of Psycholinguistic Research, 2*: 267-78.

Delattre, A.M., Liberman, A.M. & Cooper, F.S. (1955). Acoustic loci and transitional cues for consonants. *The Journal of the Acoustic Society of America, 27*, 769-73.

Dell, Gary S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93*, 282-321.

Dell, Gary S. & O'Seaghdha, Padraig G. (1992). Stages of lexical access in language production. *Cognition, 42*, 287-324.

Dell, Gary S., Martin, Nadine, Saffran, Eleanor M., Schwartz, Myrna F. & Gagnon, Deborah A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*, 801-38.

Dietrich, Rainer, Klein, Wolfgang, & Noyau, Colette. (1995). *The Acquisition of Temporality in a Second Language*. Amsterdam: Benjamins.

Doughty, Catharine. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition, 13*, 431-69.

Dowty, David. (1979). *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

Dulay, Heidi, & Burt, Marina K. (1973). Should we teach children syntax? *Language Learning, 24*, 245-58.

Dulay, Heidi, & Burt, Marina K. (1974). Natural sequences in child second language acquisition. *Language Learning, 24*, 37-53.

Dulay, Heidi., Burt, Marina K., & Krashen, Stephen D. (1982). *Language Two*. Oxford: Oxford University Press.

Durst-Andersen, Per. (1992). *Mental Grammar: Russian Aspect and Related Issues*. Columbus, Ohio: Slavica.

Eckman, Fred R. (1996). A Functional-Typological approach to second language acquisition theory. in Ritchie, William C. & Bhatia, Tej K. (eds.), *Handbook of Second Language Acquisition*, pp. 195-211. San Diego: Academic Press.

Ehrman, Madeline E. (1996). *Understanding Second Language Learning Difficulties.* Thousand Oaks, Calif.: Sage.

Ellis, Rod. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition, 9,* 1-20.

Ellis, Rod. (1990). *Instructed Second Language Acquisition.* Oxford: Blackwell.

Epstein, Samuel David, Flynn, Suzanne & Martohardjono, Gita. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences, 19,* 677-713.

Erikson, Thomas D. & Mattson, Mark E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior, 20,* 540-52.

Eubank, Lynn (ed.). (1991). *Point Counterpoint: Universal Grammar in the Second Language.* Amsterdam: Benjamins.

Eubank, Lynn. (1994). Discussion paper—Towards an explanation for the late acquisition of agreement in L2 English. *Second Language Research, 10,* 84-93.

Feldman, Laurie, B. & Fowler, Carol A. (1987). The inflected noun system in Serbo-Croation: Lexical representation of morphological structure. *Memory & Cognition, 15,* 1-12.

Felix, Sascha W. (1987). *Cognition and Language Growth.* Dordrecht: Foris.

Felix, Sascha W. (1988). UG-generated knowledge in adult second language acquisition. In Flynn, Suzanne & O'Neil, Wayne, *Linguistic Theory in Second Language Acquisition,* pp. 277-94. Dordrecht: Kluwer.

Ferguson (1987). Repertoire universals, markedness, and second language acquisition. In Rutherford, William, E. (ed.), *Language Universals and Second Language Acquisition,* pp. 247-58. Amsterdam: Benjamins.

Fillmore, Charles C. (1986). The case for case. In Emmon Bach & Robert T. Harms (eds.), *Universals in Linguistic Theory.* New York: Holt, Rinehart & Winston, 1-88.

Flashner, Vanessa E. (1989). Transfer of aspect in the English oral narratives of native Russian speakers. In Dechert, M. & Raupauch M. (eds.), *Transfer in Language Production,* pp. 71-97. Norwood, N.J.: Ablex.

Flege, James Emil. (1991). Perception and production: The relevance of phonetic input to L2 phonological learning. In Huebner, Thom & Ferguson, Charles A. (eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theories,* pp. 249-89. Amsterdam: Benjamins.

Fodor, Jerry. (1975). *The Language of Thought.* New York: Crowell.

Foster, Pauline & Skehan, Peter. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*, 299-323.

Frazier, Lyn & Clifton, Charles, Jr. (1996). *Construal.* Cambridge, Mass.: MIT Press.

Frazier, Lyn & Friederici, Angela. (1991). On deriving the properties of agrammatic Comprehension. *Brain and Language, 40*, 51-66.

Friederici, Angela D. (1985). Levels of processing and vocabulary types: Evidence from on-line comprehension in normals and agrammatics. *Cognition, 19*, 133-66.

Friederici, Angela D. (1990). On the properties of cognitive modules. *Psychological Research, 52*, 175-80.

Friederici, Angela D., Wessels, J.M.I., Emmorey, Karen, & Bellugi, Ursula. (1992). Sensitivity to inflectional morphology in aphasia: A real-time perspective. *Brain and Language 43*, 747-63.

Garnham, Alan. (1981). Mental models as representations of text. *Memory and Cognition, 9*, 560-65.

Garrett, Merrill F. (1982). Production of speech: Observations from normal and pathological language use. In Ellis, A (ed.), *Normality and Pathology in Cognitive Functions*, pp. 19-26. New York: Academic Press.

Garrett, Merrill F. (1988). Processes in language production. In Newmeyer, Frederick J. (ed.), *Linguistics: The Cambridge Survey: III. Language: Psychological and Biological Processes*, pp. 69-96. Cambridge: Cambridge University Press.

Garrod, Simon C. & Sanford, Anthony J. (1994). Selective processing in text understanding. in Gernsbacher, Morton Ann (ed.), *Handbook of Psycholinguistics*, pp. 609-40. New York: Academic press.

Gass, Susan M. (1987). The resolution of conflicts among competing systems: A bidirectional perspective. *Applied Psycholinguistics, 8*, 19-37.

Gass, Susan M. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics 9*, 198-217.

Gass, Susan M. (1994). The reliability of second language grammaticality judgments. Tarone, Elaine E., Gass, S. M. & Cohen, Andrew D. (eds.), *Research Methodology in Second-Language Acquisition.* Hillsdale, N.J.: Erlbaum.

Gass, Susan M. (1997). *Input, Interaction and the Second Language Learner.* Mahwah, N.J.: Erlbaum.

Gass, Susan, Mackey, Alison, Alvera-Torres, María José, & Fernández-García, Marisol. (1999). The effects of task repetition on linguistic output. *Language Learning, 49*, 549-81.

Gass, Susan M. & Selinker, Larry. (1994). *Second Language Acquisition: An Introductory Course*. Hillsdale, N.J.: Erlbaum.

Giacalone Ramat, A. & Banfi, E. (1990). The acquisition of temporality: A second language perspective. *Folia Linguistica 24*,405-28.

Givón, Talmy. (1983). Topic continuity in discourse: An introduction. In Givón, T. (ed.), *Topic Continuity in Discourse: A Quantitative Cross-Language Study*, pp. 5-41. Amsterdam: Benjamins.

Givón, Talmy. (1984). *Syntax: A Functional Typological Introduction, Vol. I.* Amsterdam: Benjamins

Givón, Talmy. (1990). *Syntax: A Functional Typological Introduction, Vol. II.* Amsterdam: Benjamins.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26*(1), 69-83.

Glenberg, Arthur M., Kruley, Peter, & Langston, William E. (1994). Analogical processes in comprehension: Simulation of a mental model. In Morton Ann Gernsbacher (ed.), *Handbook of Psycholinguistics*. New York: Academic Press.

Gorelov, I.N. & Sedov, K.F. (1997). *Osnovy Psixolingvistiki*. Moscow: Labirint.

Green, Peter S. & Hecht, Karlheinz. (1992). Implicit and explicit grammar: an empirical study. *Applied Linguistics, 13*, 77-97.

Green, Peter S. & Hecht, Karlheinz. (1993). Pupil self-correction in oral communication in English as a foreign language. *System, 21*, 151-63.

Greenberg, Joseph. (1966). *Language Universals, with Specific Reference to Feature Hierarchies*. The Hague: Mouton.

Gregg, Kevin. (1996). The logical and developmental problems of second language acquisition. In Ritchie, William C. & Bhatia, Tej K. (eds.), *Handbook of Second Language Acquisition*, pp. 50-81. San Diego: Academic Press.

Günther, Hartmut. 1988. Oblique word forms in visual word recognition. *Linguistics 26*, 583-600.

Gupta, Prahlad & MacWhinney, Brian. (1995). Is the articulatory loop articulatory or auditory? Reexamining the effects of concurrent articulation on immediate serial recall. *Journal of Memory and Language, 34*, 63-88.

Gupta, Prahlad & MacWhinney, Brian. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267-333.

Gvozdev, A.N. (1949). *Formirovanie u Reb'onka Grammatcheskogo Stroja Russkogo Jazyka*, parts I and II. Moscow: Academy of Pedagogical Sciences, RSFSR.

Heine, Bernd, Claudi, Ulrike, & Hünnemeyer, Friederike. (1991). *Grammaticalization: A Conceptual Framework.* Chicago: University of Chicago Press.

Hockett, Charles F. (1947). Problems of morphemic analysis. *Language 23*, 321-43.

Hockett, Charles R. (1954). Two models of grammatical description. *Word 10*, 229-42.

Hopper, Paul J. (1979). Aspect and foregrounding in discourse. In Givón, Talmy (ed.), *Syntax and semantics: Discourse and syntax*, pp. 213-241. New York: Academic Press.

Hopper, Paul J. (1982). Aspect: Between discourse and grammar. In Hopper, Paul J. (ed.) *Tense-Aspect: Between Semantics and Pragmatics*, pp. 2-18. Amsterdam: Benjamins.

Hopper, Paul J. & Thompson, Sandra A. (1980). Transitivity in grammar and discourse. *Language 56*, 251-99.

Hopper, Paul J. & Traugott, Elizabeth C. (1993). *Grammaticalization.* Cambridge University Press.

Housen, Alex. (1994). Tense and aspect in second language acquisition: The Dutch interlanguage of a native speaker of English. In Vet, Co & Vetters, Carl (eds.), *Tense and Aspect in Discourse*, pp. 257-91. Berlin: Mouton de Gruyter.

Hulstijn, Jan H. (1990). A comparison between the information-processing and the Analysis/Control approaches to language learning. *Applied Linguistic, 11*, 30-45.

Hulstijn, Jan H. & Hulstijn Wouter. (1984). Grammatical Errors as a Function of Processing Constraints and Explicit Knowledge. *Language Learning, 34*, 23-43.

Jarvella, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior, 10*, 409-16.

Jerema, Gonia & Kehayia, Eva. (1992). Impairment of inflectional morphology and lexical storage. *Brain and Language, 43*, 541-64.

Jiang, Nan. (1998). Lexical representation and development in a second language. *Applied Linguistics, 21*, 47-77.

Johnson, Jacqueline S. & Newport, Elissa L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (1989). Mental models. In M. I. Posner (Eds.), *Foundations of Cognitive Science*, pp. 469-99. Cambridge, Mass.: MIT Press.

Johnson-Laird, P. N., & Garnham, A. (1980). Descriptions and discourse models. *Linguistics and Philosophy, 3*, 371-93.

Johnson-Laird, Philip N. (1981). Mental models of meaning. In Joshi, Aravind K., Sag, Ivan & Nash-Webber, Bonnie, (eds.), *Elements of Discourse Understanding*, pp. 106-126. Cambridge: Cambridge University Press.

Juffs, Alan. (1998a). Main verb versus reduced relative clause ambiguity resolution in L2 sentence processing. *Language Learning 48*, 107-47.

Juffs, Alan (1998b). Some effects of first language argument structure and morphosyntax on second language sentence processing. *Second Language Research, 14*, 406-24.

Juffs, Allan, & Harrington, Michael. (1995). Parsing effects in L2 sentence processing: Subject and object asymmetries in wh- extraction. *Studies in Second Language Acquisition, 17*, 483-516.

Juffs, Allan & Harrington, Michael. (1996). Garden path sentences and error data in second language sentence processing research. *Language Learning, 46*, 286-324.

Keenan, Janice M., MacWhinney, Brian, & Mayhew, Deborah. (1984). Pragmatics in memory: A study of natural conversation. *Journal of Verbal Learning and Verbal Behavior, 16*, 549-60.

Kempe, Vera & MacWhinney, Brian. (1996). The crosslinguistic assessment of foreign language vocabulary learning. *Applied Psycholinguistics 17*, 149-83.

Kempe, Vera & MacWhinney, Brian. (1998). The acquisition of case marking by adult learners of Russian and German. *Studies in Second Language Acquisition, 20*, 534-58.

Kempe, Vera & MacWhinney, Brian. (1999). Processing of morphological and semantic cues in Russian and German. *Language and Cognitive Processes, 14*, 129-71.

Kempen, Gerard & Hoenkamp, Edward. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science 11*, 201-58.

Kess, Joseph F. (1992). *Psycholinguistics: Psychology, Linguistics and the Study of Natural Language*. Amsterdam: Benjamins.

Kilborn, Kerry. (1991). Selective impairment of grammatical morphology due to induced stress in normal listeners: Implications for aphasia. *Brain and Language, 41*, 275-88.

Kilborn, Kerry & Ito Takehiko. (1989). Sentence processing strategies in adult bilinguals. In MacWhinney, Brian & Bates, Elizabeth (eds.), *The Crosslinguistic study of sentence processing.* Cambridge: Cambridge University Press.

Kilby, David. (1986). The instrumental in Russian: On establishing a consensus. In Brecht, Richard D. & Levine, James S. (eds.), *Case in Slavic*, 333-37. Columbus: Slavica.

Kintsch, Walter & Bates, Elizabeth. (1977). Recognition memory for statements from a classroom lecture. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 150-59.

Klein, Wolfgang. (1986). *Second Language Acquisition.* Cambridge: Cambridge University Press.

Klein, Wolfgang. (1995). A time-relational analysis of Russian aspect. *Language, 71,* 669-95.

Koctomarov, V.G. & Mitrofanova, O.D. (1988). *Metodicheskoe Rukovodstvo dl'a Prepodovatelej Russkogo Jazyka Inostrancam.* Moscow: Russkij Jazyk.

Kosslyn, Stephen M. (1994). *Imagery and Brain.* Cambridge, Mass.: MIT Press.

Krashen, Stephen D. (1977). Some issues relating to the Monitor Model. In (Brown, H. Douglas, Yorio, Carlos, & Crymes, R. (eds.), *On TESOL '77,* pp. 144-58. Washington, D.C.: TESOL.

Krashen, Stephen D. (1981). *Second Language Acquisition and Second Language Learning.* Oxford: Pergamon.

Krashen, Stephen. (1982). *Principles and Practices in Second Language Acquisition.* Oxford: Pergamon.

Krashen, Stephen D. (1985). *The Input Hypothesis.* London: Longman.

Krashen, Stephen D., Butler, J., Birnbaum, R. & Robertson, J. (1978). Two studies in language acquisition and language learning. *ITL: Review of Applied Linguistics, 39-40,* 73-92.

Krashen, Stephen, Sferlazza, Victoria, Feldman, Lorna, & Fathman, Ann K. (1976). Adult performance on the SLOPE test: More evidence for a natural sequence in adult second language acquisition. *Language Learning, 26,* 145-51.

Kroll, Judith F. & Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In Gruneberg Michael M., Morris, Peter Edwin & Sykes, Robert N. (eds.), *Practical Aspects of Memory, vol. 2,* pp. 389-95. London: Wiley.

Kroll, Judith R. & de Groot, Annette M.B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In de Groot, A.M.B.

& Kroll, J.R. (eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives,* pp. 169-99.

Lakoff, George & Johnson, Mark. (1980). *Metaphors We Live By.* Chicago: University of Chicago Press.

Langacker, Ronald W. (1991). *Foundations of Cognitive Grammar. Vol. 2: Descriptive Application.* Stanford: Stanford University Press.

Larsen-Freeman, Diane, & Long, Michael H. (1991). *An Introduction to Second Language Acquisition Research.* London: Longman.

Leather, Jonathan & Allan, James. (1996). Second language speech and the influence of the first language. In Ritchie, William C. & Bhatia, Tej K. (eds.), *Handbook of Second Language Acquisition* pp. 269-316. San Diego: Academic Press.

Lee, Dong Jae. (1981). Interpretation of morpheme rank ordering in L2 research. In P. Dale & D. Ingram (Eds.), *Child language: An international perspective,* pp. 261-71. Baltimore, Md.: University Park Press.

Levelt, Willem J.M. (1978). Skill theory and language teaching. *Studies in Second Language Acquisition, 1,* 53-70.

Levelt, Willem. (1989). *Speaking: From Intention to Articulation.* Cambridge, Mass.: MIT Press.

Libben, Gary. (2000). Representation and processing in the second language lexicon: The homogeneity hypothesis. In Archibald, John (ed.), *Second Language Acquisition and Linguistic Theory.* Oxford: Blackwell.

Liberman, Alan M. (1970). The grammars of speech and language. *Cognitive Psychology, 1,* 301-23.

Liberman, Alan M., Cooper, F.S., Shankweiler, D.P. & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74,* 431-61.

Liberman, Alan M., Delattre, P.C., & Cooper, F.S. (1952). The role of selected stimulus variables in the perception of unvoiced stop consonants. *The American Journal of Psychology, 65,* 497-516.

Lieber, Rochelle. (1992). *Deconstructing Morphology.* Chicago: University of Chicago Press.

Lightbown, Patsy M. (1998). The importance of timing in focus on form. In Doughty, Catherine & Williams, Jessica (eds.), *Focus on Form in Classroom Second Language Acquisition.* Cambridge: Cambridge University Press.

Lightbown, Patsy M. & Spada, Nina. (1990). Focus on form and corrective feedback in communicative language teaching. *Studies in Second Language Acquisition, 12,* 429-48.

Lima, Susan D. Morphological analysis in sentence reading. *Journal of Memory and Language, 26,* 84-99.

Liskin-Gasparro, J. (1987). *Testing and teaching for oral proficiency.* Boston: Heinle & Heinle.

Lotto, Lorella & de Groot, Annette M.B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning, 48,* 31-69.

Lukatela, G., Gligorijevič, B., Kostič, A. & Turvey, M.T. (1980). Representation of inflected nouns in the internal lexicon. *Memory and Cognition, 8,* 415-23.

Lukatela, G., Mandič, Z. Gligoriavič, B., Kostič, A., Savič, M. & Turvey, M.T. (1978). Lexical decision for inflected nouns. *Language and Speech, 21,* 166-73.

Luria, A.R. (1946). *O patologii grammaticheskix operatsi.* Moscow: ANP RSFSR.

Lyster, Roy. (1998). Negotiation of form, recasts, and explicit correction in relation to error types in learner repair in immersion classrooms. *Language Learning, 48,* 183-218.

Lyster, Roy & Ranta, Leila. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 19,* 37-66.

MacDonald, Maryellen. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes, 9,* 257-201.

MacKain, K.S., Best, C.T. & Strange, W. (1981). Categorical perception of /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics,* 2, 655-69.

MacWhinney, Brian. (1997). Second language acquisition and the Competition Model. In de Groot, Annette M.B., & Kroll, Judith F. (eds.), *Tutorials in Bilingualism,* pp. 113-42. Mahwah, N.J.: Erlbaum.

Mani, K., & Johnson-Laird, Philip N. (1982). The mental representation of spatial descriptions. *Memory and Cognition, 10,* 181-87.

Marslen-Wilson, William D. & Tyler, Lorraine Komisarjevky. (1980). The temporal structure of spoken language understanding. *Cognition, 8,* 1-71.

Marslen-Wilson, William D. (1975). Sentence perception as an interactive parallel process. *Science 152,* 226-8.

Marslen-Wilson, William D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25,* 71-102.

Marslen-Wilson, William D. & Welsh, A. (1978). Processing interaction and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*, 29-63.

Marslen-Wilson, William D. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature, 244*, 522-23.

Masny, Diana & d'Anglejan, Alison. (1985). Language, cognition and second language grammaticality judgments. *Journal of Psycholinguistic Research, 14*, 175-97.

Matthews, Peter Hugoe. (1991). *Morphology (2nd ed.)*. Cambridge: Cambridge University Press.

McCarthy, John J. (1981.) A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry 1*, 373-418.

McLaughlin, Barry. (1978). The Monitor Model: Some methodological considerations. *Language Learning, 28*, 309-32.

McLaughlin, Barry. (1987). *Theories of Second Language Learning*. London: Edward Arnold.

McLaughlin, Barry. (1990). Restructuring. *Applied Linguistics, 11*, 113-28.

McLaughlin, Barry & Heredia, Roberto. (1996). Information-processing approaches to research on second language acquisition and use. In Ritchie, William C. & Bhatia, Tej K. (eds.), *Handbook of Second Language Acquisition*, pp. 213-28.

McLauglin, Barry, Rossman, Tammi, & McLeod, Beverly. (1983). Second language learning: An information-processing perspective. *Language Learning, 33*,135-58.

McRae, Ken, Ferretti, Todd R., & Amyote, Liane. (1997). Thematic roles as verb-specific concepts. *Language & Cognitive Processes, 12*, 137-76.

McRae, Ken, Spivey-Knowlton, Michael J, & Tanenhaus, Michael K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38*, 283-312.

Meisel, Jürgen M. (1987). Reference to past events and actions in the development of natural second language acquisition. In Andersen, Roger W. (ed.), *Cross-linguistic Series on Second Language Research*, pp. 206-24. Cambridge: Newbury House.

Meisel, Jürgen M., Clahsen, Harald, & Pienemann, Manfred. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition 3*, 109-35.

Mitchell, Rosamond & Myles, Florence. (1998). *Second Language Learning Theories*. New York: Arnold.

Morrow, Daniel G. (1985). Prepositions and verb aspect in narrative understanding. *Journal of Memory and Language 24*, 390-404.

Morrow, Daniel G. (1986). Grammatical morphemes and conceptual structure in discourse processing. *Cognitive Science 10*, 423-55.

Morrow, Daniel G. (1990). Spatial models, prepositions and verb aspect markers. *Discourse Processes 13*, 441-69.

Nelson, Andrea. (1998). *The Ontogenesis of Grammar: An Empirical Study of the Nature, Place and Role of Formulaic Language in Russian Foreign Language Learning*. Ph.D. dissertation, Bryn Mawr College.

Nida, Eugene A. (1948). The identification of morphemes. *Language 24*, 441-41.

Nida, Eugene A. (1949). *Morphology: The Descriptive Analysis of Words* (2nd edition). Ann Arbor: University of Michigan Press.

Offord, Derek. (1996). *Using Russian*. Cambridge: Cambridge University Press.

Paivio, Alan. (1971). *Imagery and Verbal Processes*. New York: Holt, Rinehart and Winston.

Paradis, Michel. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism. In Nick Ellis (ed.), *Implicit and Explicit Learning of Languages*, pp. 393-419. New York: Academic Press.

Peters, Ann M., & Menn, Lisa. (1993). False starts and filler syllables: Ways to learn grammatical morphemes. *Language, 69*, 742-77.

Pica, Teresa. (1988). Morpheme data analysis in second language acquisition research: Renewing an old debate and raising new issues. *I.T.L., Review of Applied Linguistics, 79-80*, pp. 77-112.

Pienemann, Manfred. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition, 6*, 186-214.

Pienemann, Manfred. (1999). *Language Processing and Second Language Development: Processability Theory*. Amsterdam: Benjamins.

Pollock, Y.-I. (1989). Verb Movement, Universal Grammar and the structure of IP. *Linguistic Inquiry 20*, 365-434.

Potter, Mary C., So, Kwok-Fai, Von Eckardt, Barbara & Feldman, Laurie B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*, 23-38.

Pupynin, Ju. A. Formirovanie vidovremennoj sistemy v detskoy rechi. In Chertkova, M. Ju. (ed.) *Trudy Aspektologicheskogo Seminara Filologicheskogo Fakul'teta MGU Im. M.V. Lomonosova, Tom 3*, pp. 157-69. Moscow: Moscow University Press.

Putnam, Hiliary. (1975). The meaning of "meaning". In Kr. Gunderson (ed.), *Language, Mind and Knowledge*, 131-93. Minneapolis: University of Minnesota Press.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior, 22*, 358-74.

Rifkin, Benjamin. (1995). Error gravity in learners' spoken Russian: A preliminary study. *Modern Language Journal, 79*, 477-90.

Ringbom, Håkan. (1992). On L1 transfer in L2 comprehension and L2 production. *Language Learning, 42*, 85-112.

Robison, Richard E. (1990). The primacy of aspect: Aspectual marking in English interlanguage. *Studies in Second Language Acquisition 12*, 315-30.

Romjin, Elizabeth & Seely, Contee. (1988). *Live Action English*. Old Tappan, N.J.: Alemany Press.

Rozental', D.E., Golub, I B. & Telenkova, M.A. (1995). *Sovremennyj Russkij Jazyk*. Moscow: International Relations.

Rubin, Joan. (1994). A review of second language listening comprehension research. *Modern Language Journal, 78*, 199-221.

Rumelhart, David E. & McClelland, James L. (1986). On learning the past tense of English verbs. In Rumelhart, D.E. & McClelland, J.L. (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, pp. 216-71. Cambridge, Mass.: MIT Press.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics, 2*, 437-42.

Samuel, Arthur B. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General, 110*, 474-94.

Sanford, Anthony J. & Garrod, Simon C. (1981). *Understanding Written Language*. Chichester: Wiley.

Schachter, Jacqueline. (1990.) On the issue of completeness in second language acquisition. *Second Language Research, 6*, 93-124.

Schachter, Jacquelyn. (1998). Recent research in language learning studies. *Language Learning, 48*, 557-83.

Schmidt, Richard W. (1990). The role of consciousness in second language learning. *Applied Linguistics 11*, 129-58.

Schmidt, Richard W. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*, 357-85.

Schmidt, Richard W. & Frota, Sylvia. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese, in Day, R. (ed.), *Talking to Learn: Conversation in Second Language Acquisition*, pp. 237-322. Rowley, Mass.: Newbury House.

Schneider, Walter & Shiffrin, Richard M. (1977). Controlled and automatic human information processing: I. Detection, Search and Attention. *Psychological Review, 84*, 1-66.

Schwartz, Bonnie D. (1993). On explicit and negative data effecting and affecting *compentence* and *linguistic behavior. Studies in Second Language Acquisition, 15*, 147-63.

Schwartz, Bonnie D. & Sprouse, Rex A. (1996). L2 cognitive states and Full Transfer/Full Access Model. *Second Language Research, 12*, 40-72.

Selinker, Larry. (1972). Interlanguage. *International Review of Applied Linguistics 10*, 209-31.

Sharwood-Smith, Michael. (1986). Comprehension versus acquisition: Two ways of processing input. *Applied Linguistics, 7*, 239-55.

Sharwood-Smith, Michael. (1996). Metalinguistic ability and primary linguistic data. *Behavioral and Brain Sciences, 19*, 740-741.

Shiffrin, Richard M. & Schneider, Walter. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*, 127-90.

Shillock, Richard. (1990). Lexical hypotheses in continuous speech. In Altmann, Gerry T.M. (ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*, Pp. 24-49. Cambridge, Mass.: MIT Press.

Shirai, Yasuhiro & Kurono, Atsuko. (1998). The acquisition of tense-aspect marking in Japanese as a second language. *Language Learning 48*, 244-47.

Shvedova, H. Ju., Arut'unova, N.D., Bondarko, A.V., Ivanov, V.V., Lopatin, V.V., Uluxanov, I.S., & Filin, F.P. (1982). *Russkaja Grammatika (vols. 1 and 2).* Moscow: Russian Language Institute, USSR Academy of Sciences.

Skehan, Peter. (1998). *A Cognitive Approach to Language Learning.* Oxford: Oxford University Press.

Slobin, Dan I. (1973). Cognitive prerequisites for the development of grammar. In Ferguson, Charles A. & Slobin, Dan I. (eds.), *Studies of child language development*, pp. 175-208. New York: Holt, Rinehart & Winston.

Slobin, Dan Isaac. (1966a). Abstracts of Soviet child language studies. In Smith, Frank & Miller, George A. (eds.) The Genesis of Language: A Psycholinguistic Approach, pp. 363-86. Cambridge, Mass.: MIT Press.

Slobin, Dan Isaac. (1966b). The acquisition of Russian as a native language. In Smith, Frank & Miller, George A. (eds.) *The Genesis of Language: A Psycholinguistic Approach*, pp. 129-48. Cambridge, Mass.: MIT Press.

Slobin, Dan Isaac. (1982a). Universal and particular in the acquisition of language. In Wanner, E. & Gleitman, Lila R. (eds.), *Language Acquisition: The State of the Art*, pp. 128-70. Cambridge: Cambridge University Press.

Slobin, Dan Isaac. (1982b). The origins of grammatical coding of events. In Syntax & Semantics 15. New York: Academic Press.

Slobin, Dan Isaac. (1985). Crosslinguistic evidence for the Language-Making Capacity. In Slobin, D.I. (ed.), *The Crosslinguistic Study of Language Acquisition, vol. 2: Theoretical Issues*, pp. 1157-1256.

Smith, Stan & Bates, Elizabeth. (1987). Accessibility of case and gender contrasts for agent-object assignment in Broca's aphasics and fluent anomics. *Brain and Language 30*, 8-32.

Sorace, Antonella. (1996). The use of acceptability judgmnets in second language acquisition research. In Ritchie, William C. & Bhatia, Tej K. (eds.), *Handbook of Second Language Acquisition*, pp. 375-409. San Diego: Academic Press.

Spada, Nina. (1997). Form-focussed instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching, 30*, 73-87.

Spada, Nina & Lightbown, Patsy M. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition, 15*, 205-24.

Sperber, Dan. & Wilson, Diedre. (1986). *Relevance: Communication and Cognition.* Cambridge: Cambridge University Press.

Stemberger, Joseph Paul. (1998). Morphology in language production with special reference to connectionism. In Spencer, Andrew & Zwicky, Arnold M. (eds.), *The Handbook of Morphology*. London: Blackwell.

Stoll, Sabine. (1998). Acquisition of Russian Aspect. *First Language 18*, 351-77).

Stowe, Laurie A. (1989). Thematic structures and sentence comprehension. In Carlson, Greg & Tannenhaus, Michael (eds.), *Linguistic Structure in Language Processing*, pp. 319-57. Dordrecht: Kluwer.

Stump, Gregory. (1993). On rules of referral. *Language, 69*, 449-79.

Stump, Gregory. (1998). Inflection. In Spencer, Andrew & Zwicky, Arnold (eds.), *The Handbook of Morphology*, pp. 13-43. Oxford: Blackwell.

Swain, Merrill. (1985). Communicative competence: Some roles for comprehensible input and comprehensible output in its development. In Gass, Susan M. & Madden, Carolyn G. (eds.), *Input in Second Language Acquisition*, pp. 235-53. Cambridge, Mass.: Newbury House.

Swinney, David A. (1979). Lexical access during sentence comprehension; (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645-59.

Taft, Marcus. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition, 7*, 263-72.

Taft, Marcus & Forster, Keith I. (1976), Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14*, 638-47.

Takahashi, Noriko, & Frauman-Prickel. (1985). *Action English Pictures: Activities for Total Physical Response.* Englewood Cliffs, N.J.: Prentice-Hall.

Tanenhaus, Michael, K. & Trueswell, John C. (1995). Sentence comprehension. In Miller, J.L. & Eimas, P.D. (eds.), *Handbook of Perception and Cognition, Vol. 11: Speech, Language and Communication*, pp. 217-62. San Diego: Academic Press.

Tarone, Elaine E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning, 35*, 373-403.

Tartter, Vivien C. (1998). *Language and its Normal Processing.* Thousand Oaks, Calif.: Sage.

Thomson, Greg & Zawaydeh, Bushra Adnan. (1996). A search for inflectional priming reveals an effect of discourse type on the lexical access of inflected verbs. *Work Papers of the Summer Institute of Linguistics, University of North Dakota, 40*, 111-25.

Thomson, Greg. (1995). Comprehensibility and processing demands as factors in the order of acquisition of *–ing* and *–ed*: A search for evidence. *University of Alberta Papers in Experimental and Theoretical Linguistics, 3*, 68-86.

Thomson, Greg. (1996). Hierarchical dependencies in comprehension abilities and the evolution of proficiency. Paper presented at the meeting of the Canadian Association of Applied Linguistics.

Tiersma, P. (1982). Local and general markedness. *Language, 58*, 832-49.

Timberlake, Alan. (1982). Invariance and the syntax of Russian aspect. In Hopper, Paul, J. (ed.), *Tense-Aspect: Between Semantics and Pragmatics*, pp. 305-331. Amsterdam: Benjamins.

Timberlake, Alan. (1985). Reichenbach and Russian Aspect. In Flier, M.S. and Timberlake, A. (eds.), *The Scope of Slavic Aspect*, pp. 153-68.

Towel, R. & Hawkings, R. (1994). *Approaches to Second Language Acquisition.* Clevendon: Multilingual Matters.

Trahey, M. & White, Lydia. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition, 15*, 181-204.

Trueswell, John C., Tanenhaus, Michael K. & Garnsey, Susan M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution.. *Journal of Memory and Language, 33*, 285-318.

Trueswell, John C., Tanenhaus, Michael K., & Kello, Christopher. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19*, 528-53.

Truscott, John. (1998). Noticing in second language acquisition: A critical review. *Second Language Research, 14*, 103-35.

Tunmer, William E. & Herriman, Michael I. (1984). The development of metalinguistic awareness: A conceptual overview. In Tunmer, W., Pratt, C. & Herriman, M. (eds.), *Metalinguistic awareness in children*, pp. 12-35. New York: Springer-Verlag.

Vainikka, Anne, & Young-Scholten, Martha. (1998). Morphosyntactic triggers in adult SLA. In Beck, Maria-Luise (ed.), *Morphology and its Interfaces in Second Language Knowledge*. Amsterdam: Benjamins.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

VanPatten, Bill. (1996). *Input Processing and Grammar Instruction: Theory and Research.* Norwood, N.J.: Ablex.

VanPatten, Bill. (2000). Thirty years of input (or intake, the neglected sibling). In Anderson, M., Klee, C. Morris, F, Tarone, E. & Swierzbin, B. (eds.), *Interaction of Social and Cognitive Factors in SLA: Selected Proceedings of the 1999 Second Language Research Forum*. Somerville, Mass.: Cascadilla Press.

VanPatten, Bill & Cadierno, Teresa. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15*, 225-41.

VanPatten, Bill & Sanz, Cristina. (1995). From input to output: Processing instruction and communicative tasks. In Eckman, Fred R., Highland, D., Lee, Peter W., Mileham, Jean & Weber, Rita Rutowski, (eds.), *Second Language Acquisition Theory and Pedagogy*. Mahwah, N.J.: Erlbaum.

VanPatten, Bill. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 12*, 287-301.

Véronique, Daniel. (1987). Reference to past events and actions in narratives in L2: Insights from North African learner's French. In Pfaff, C.W. (ed.), *First and Second Language Acquisition*, pp. 252-72. Rowley, Mass.: Newbury House.

von Stutterheim, Christiane. (1991). Narrative and description: Temporal reference in second language acquisition. In Huebner, Thom & Ferguson, Charles A. (eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theories*, pp. 385-403. Amsterdam: Benjamins.

Wade, Terrence. (1992). *A Comprehensive Russian Grammar.* London: Blackwell.

Wagner-Gough, K., & Hatch, E. (1975). The importance of input in second language acquisition studies. *Language Learning, 25*, 297-308.

Warren, Richard M. 1970. Perceptual restoration of missing speech sounds. *Science, 167*, 392-93.

Warren, Richard M. & Sherman, Gary L. (1974). Phonetic restorations based on subsequent context. *Perception and Psychophysics, 16*, 150-56.

Wason, P. & Reich, S.S. (1979). A verbal illusion. *Quarterly Journal of Experimental Psychology, 31*: 591-97.

Weist, Richard M. (1986). Tense and aspect. In Fletcher, Paul & Garman, Michael (eds.), *Language Acquisition*, 2nd edition, pp. 356-74. Cambridge: Cambridge University Press.

Weist, Richard M., Wysocka, H., Witkowska-Stadnik, K. Buczowska, E., & Konieszna, E. (1984). The defective tense hypothesis: on the emergence of tense and aspect in child Polish. *Journal of Child Language 11*, 347-74.

Werker, J.F. (1993). Developmental changes in cross-language speech perception: Implications for cognitive models of speech processing. In G.T.M. Altmann & R.C. Shillcock (eds.), *Cognitive Models of Speech Processing: The Second Sperlonga Workshop*, pp. 57-78. Hillsdale, N.J.: Erlbaum.

White, Lydia. (1989). The adjacency condition on case assignment: do L2 learners observe the Subset Principle? In Gass, Susan M. & Schachter, Jacquelyn (eds.), *Linguistic Perspectives on Second Language Acquisition.* Cambridge: Cambridge University Press.

White, Lydia. (1991a). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research, 7*, 133-61.

White, Lydia. (1991b). The verb-movement parameter in second language acquisition. *Language Acquisition, 1*, 273-86.

White, Lydia. (1996). Universal Grammar and second language acquisition. In Ritchie, W.C. & Bhatia, Tej K. (eds.), *Handbook of Second Language Acquisition*, pp. 85-120. New York: Academic Press.

White, Lydia, Spada, Nina, Lightbown, Patsy M. & Ranta, Leila. (1991). Input enhancement and L2 question formation. *Applied Linguistics, 12,* 416-32.

Wilson, S. G., Rinck, M., McNamara, T. P., Bower, G. H., & Morrow, D. G. (1993). Mental models and narrative comprehension:. *Journal of Memory and Language, 32,* 141-54.

Wittgenstein, Ludwig. (1958). *Philosophical Investigations.* 2nd ed. Trans. Anscombe, G.E.M. Oxford: Blackwell.

Wode, Henning, Bahns, H. Bedey, & Frank, W. (1978). Developmental sequence: An alternative approach to morpheme order. *Language Learning, 28,* 175-185.

Yang, Lynne R. & Givón, Talmy. (1997). Benefits and drawbacks of controlled laboratory studies of second language acquisition: The Keck second language learning project. *Studies in Second Language Acquisition, 19,* 173-93.

Zakharova, A.V. (1973 [1958]). Acquisition of forms of grammatical case by preschool children. In Slobin, Dan I. & Ferguson, Charles A. (eds.), *Studies of Child Language Development,* pp. 281-92.

Zobl, Helmut & Liceras, Juana (1994). Review article: Functional categories and acquisition orders. *Language Learning 44,* 159-80.

Zwicky, Arnold. (1985). How to describe inflection. *Proceedings of the Eighth Annual Meeting of the Berkeley Linguistics Society,* pp. 372-386.

Zwitserlood, Pienie. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32,* 25-64.

Appendix: Texts from Experiments 1 and 2

Experiment 1, Text 1

(Erroneous word-forms in bold-face; correct targets in parentheses)

1. У одного пожилого человека были грязные руки. Поэтому он хотел руки вымыть, и пошёл к раковине. Рядом с раковиной на **вешалку** (вешалке) висело полотенце.
2. Пожилой человек открыл кран.
3. Потом он взял мыло из мыльницы. Мыльница осталась пустой.
4. Потом **человеком** (человек) стал мыть руки под краном.
5. После того как намылил руки **мыле** (мылом), он положил его в мыльницу.
6. Потом он продолжал мыть руки.
7. Теперь у человека руки чистые. Он сразу **закрывал** (закрыл) кран.
8. Вода капала с рук. Поэтому он снял полотенце с вешалки.
9. Наконец человек вытер руки **полотенце** (полотенцем), и вода перестала капать.
10. Теперь у человека руки чистые и сухие. Он повесил полотенце на место.

Translation: 1) There was an elderly person with dirty hands. And so he wanted to wash his hands. So he went to the sink. 2) The elderly person turned on the tap. 3) Next he took the soap from the soap dish and the soap dish lay empty. 4) Next the person started washing his hands under the tap. 5) After he rubbed soap on his hands, he lay the soap back in the soap dish. 6) Next he continued washing his hands. 7) Then the person had clean hands. He at once turned off the tap. 8) Water dripped from his hands. So he took the towel from the rack. 9) Finally the person wiped his hands with the towel, and the water stopped dripping. 10) Then the person's hands were clean and dry. He hung the towel back up.

Text 2

1. Десять часов вечера. Уютная комната. Пожилой человек сидит в кресле и смотрит телевизор.
2. На экране телевизора появилась надпись „Конец.
3. Человек очень устал сидеть в **кресло** (кресле) перед телевизором.
4. Он встал, подошел к телевизору и выключил его.
5. Потом он потянулся, сладко зевнул.
6. И пошел в спальню, где **постояли** (стояли) кровать, столик и стул.
7. Там человек подошел к кровати, включил лампу на столике, чтобы лучше видеть.

8. Когда он сидел на кровати, он снял тапочки и носки.
9. Потом он поднялся с кровати, снял рубашку.
10. Теперь его рубашка и **брюках** (брюки) висят на стуле, а сам он в пижаме.
11. После этого человек приготовил постель, снял покрывало с кровати.
12. Потом он выключил лампу, и в спальне стало темно.
13. Он сел в кровати.
14. Когда он лежал в **кроватью** (кровати), он поправил подушку под головой.
15. Натянул на себя одеяло.
16. Потом он **закрывал** (закрыл) глаза и уснул.
17. Он лежит на подушке под одеялом и крепко спит. Во сне он видит свою семью.

Translation: It is 10:00 in the evening. There is a cozy room. An elderly person is sitting in an armchair and watching television. 2) On the television screen appeared the words "The End". 3). The person is very tired of sitting in the armchair in front of the television. 4) He stood up, walked to the television, and turned it off. 5) Next he stretched, and had a good yawn. 6) Next he walked into the bedroom, where there stood a bed, a small table and a chair. 7) There the person walked to the bed and turned on the lamp on the table, in order to see better. 8) When he sat on the bed, he pulled off his shoes and socks. 9) Next he got up from the bed, and took off his shirt. 10) Then his shirt and pants are hanging on the chair, and he himself is wearing pajamas. 11) After this, the person prepared the bedding, and pulled back the blanket from the bed. 12) Next he turned off the lamp, and the room became dark. 13) He sat in the bed. 14) As he lay in the bed, he adjusted the pillow under his head. 15) He pulled the covers over himself. 16) Next he closed his eyes and fell asleep. 17) He lies on the pillow under the blanket and sleeps soundly. In his dream he sees his family.

Experiment 2, Text 1

1. У одного пожилого мужчины были грязные руки. Поэтому он хотел руки вымыть и пошёл к **раковину** (раковине).
2. Потом этот человек открыл кран и вода стала течь из него.
3. Потом **человека** (человек) взял мыло из мыльницы. Мыльница осталась пустой.
4. Тогда он протянул руки под кран и стал их мыть.
5. После того как он намылил руки **мыле** (мылом), он положил его в мыльницу.
6. Потом он продолжал мыть руки.
7. Теперь у мужчины руки чистые. Он закрыл кран.
8. Он снял полотенце с вешалки.
9. Он вытер руки полотенцем и вода **перестал** (перестала) капать с рук.
10. Человек повесил полотенце на место. Теперь у него руки чистые и сухие.

Translation: 1) There was an elderly person with dirty hands. So he wanted to wash his hands. So he went to the sink. 2) Next this elderly person turned on the tap, and water started flowing from it. 3) Next he took the soap from the soap dish. The soap dish lay empty. 4) Then he put his hands under the tap and started washing them. 5) After he rubbed soap on his hands, he lay the soap back in the soap dish. 6) Next he continued washing his hands. 7) Then the man had clean hands. He turned off the tap. 8) He took the towel from the rack. 9) He wiped his hands with the towel, and the water stopped dripping from his hands. 10) The person hung the towel back up. Then his hands were clean and dry.

Text 2

1. Десять часов вечера. Уютная комната. Человек **сидишь** (сидит) в кресле и смотрит телевизор.
2. Но телепрограмма подошла к концу.
3. Мужчина вздремнул в кресле перед телевизором.
4. Через некоторое время **мужчину** (мужчина) проснулся, встал, подошел к телевизору и выключил его.
5. Потом он потянулся и сладко зевнул.
6. Он пошел в спальню, где стояли кровать, лампа на столике и стул.
7. В спальне человек подошел к кровати, включил **лампа**, (лампу) чтобы лучше видеть.
8. Он сел на постель, снял тапочки и носки и остался в рубашке и брюках.
9. Потом он поднялся, снял **рубашка** (рубашку) и остался в брюках.
10. Теперь его рубашка и брюки висят на стуле, а сам он в пижаме.
11. После этого человек приготовил постель, снял покрывало с кровати.
12. Потом он выключил лампу и в спальне стало темно.
13. Он сел на постель и накрыл ноги одеялом.
14. Когда он лежал в **кроватью** (кровати) он поправил подушку под головой.
15. Теперь он натянул на себя одеяло.
16. Потом он **закрывал** (закрыл) глаза и уснул на спине.
17. Он повернулся набок. Он лежит на подушке под одеялом и крепко спит. Во сне он видит свою семью.

Translation: It is 10:00 in the evening. There is a cozy room. A person is sitting in an armchair and watching television. 2) But the program ended. 3). The man dozed off in the armchair in front of the television. 4) After a little while, the man awoke, stood up, walked to the television, and turned it off. 5) Next he stretched, and had a good yawn. 6) He walked into the bedroom, where there stood a bed, a small table and a chair. 7) There the person walked to the bed and turned on the lamp, in order to see better. 8) He sat on the bed, pulled off his shoes and socks, and remained in his shirt and pants. 9) Next he got up, took off his shirt, and remained in his pants. 10) Now his shirt and pants are hanging on the chair, and he himself is wearing pajamas. 11) After this, the person prepared the bedding, and pulled back the blanked from the bed. 12) Next he turned off the lamp, and the room became dark. 13) He sat on the bed and covered his feet with the blanket. 14) As he lay in the bed, he adjusted the pillow under his head. 15) He pulled the covers over himself. 16) Next he closed his eyes and fell asleep. 17) He turned on his side. He lies on the pillow under the blanket and sleeps soundly. In his dream he sees his family.