BEGINNING AND ADVANCED LEARNERS’ AWARENESS OF CORRECTIVE FEEDBACK IN THE ARABIC FOREIGN LANGUAGE CLASSROOM

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Arabic

By

Gergana Atanassova, M.A.

Washington, DC
April 20, 2012
Copyright 2012 by Gergana Atanassova
All Rights Reserved
BEGINNING AND ADVANCED LEARNERS’ AWARENESS OF CORRECTIVE FEEDBACK IN THE ARABIC FOREIGN LANGUAGE CLASSROOM

Gergana Atanassova, M.A.

Thesis Advisor: Kassem Wahba, Ph.D.

ABSTRACT

Corrective feedback as part of conversational interaction has been shown to facilitate language development (Li, 2010; Russel & Spada, 2006), but learners differ in the extent to which they benefit from it. Awareness is one cognitive process that has been proposed to play a supportive (Gass, 1997; Schmidt, 2009) to essential (Carroll, 2001) role in moderating feedback effectiveness. Language proficiency has also been suggested to influence the way learners engage with feedback (Ammar & Spada, 2006; Philp, 2003). However, the relationship between proficiency and awareness of feedback during classroom interaction has not yet been empirically addressed.

In this study, five intact Arabic classes – four beginning and one advanced – were observed and videotaped during unscripted whole-class conversational activities. Volunteers from each class (26 beginners, 5 advanced) then participated in a stimulated recall interview in order to gauge their awareness of the target and corrective intent of classroom feedback. The interview protocols were coded for presence or absence of awareness, and analyzed in relation to three characteristics of the feedback episode to which they were referring: type, linguistic target, and direction. Qualitative thematic analysis of the protocols was used to identify additional factors in learners’ awareness of feedback.
Advanced learners reported awareness of correction 79% and awareness of target 41% of the time, significantly more frequently than beginners (54% and 27%, respectively). None of the feedback characteristics investigated in the study was significantly associated with advanced learners’ awareness, but all three were significantly associated with beginning learners’ awareness. Thematic analysis of interview protocols showed that learner awareness was additionally influenced by the presence of new information in the episode, affective response to the feedback or its addressee, and external factors such as fatigue. The findings highlight the role of proficiency in feedback awareness and suggest a need for further exploration of the influence of affect on learner engagement with feedback.
ACKNOWLEDGMENTS

This dissertation would not have been possible without the generosity of many.

I am forever indebted to:

The thirty one learners of Arabic who participated in this study, and the wonderful teachers in the Georgetown Arabic program who let me into their classrooms and gave generously of their time and insights into the teaching of Arabic: Margaret Nydell, Amin Bonnah, Kassem Wahba, Hana Zabarah, Aja Ishmael, Ben Ellis, Katrien Vanpee and Hanaa Kilany. I have been extremely lucky to have them as my teachers, colleagues, mentors, friends. I owe special thanks to Hana Zabarah, who not only allowed me to become a fixture of her classrooms for almost two semesters, but also spent hours upon hours of her precious time talking with me about feedback, teaching, the classroom learning of Arabic, and life in general.

My committee – Reem Bassiouney for readily agreeing to serve as my reader on a very short notice; Terrence Potter for his astute comments and thought-provoking feedback; Kassem Wahba for stepping in as advisor at the last moment, for reading numerous drafts at lightning speed, for his unshakable optimism, and for his unwavering support and faith in my work.

The great Karin Ryding – for meeting all my research ideas with warm enthusiasm and helping me feel that my work is important and needed.

Alison Mackey – for some of the most important introductions in my research and personal life so far. Alison introduced me to research design in general and classroom research in
particular, to the interaction approach, to a standard of research that I will always strive to emulate, and to the most fantastic research group ever. This study was inspired by Alison’s work and set into motion with her generous help, feedback and encouragement. Thank you.

Katrien Vanpee and Emma Trentman for assisting with the coding and for being my friends in the land of Arabic; Maymona AlKhalil for many conversations on noticing and feedback and for being there when it mattered; RuSan Chen for help with the statistical analysis; Leslie Moore at the Ohio State University for a most inspirational course on SLA at a crucial point in my work.

Mika Hama – for being the best friend one could hope for.

My wonderful parents – for loving and supporting me always, for flying over the ocean twice to be with me and help me at the toughest points in the journey. Their pride in my work has been my inspiration.

And last but in no way least - my long-suffering husband, who endured not one or two, but nine whole years of living with a graduate student, and (mostly) graciously accepted the dust bunnies, cold dinners, missed family evenings and the emotional rollercoaster that the preparation of this manuscript turned out to be. With his help, encouragement and, eventually, increasing impatience, it is now finally done.

Many thanks.
This dissertation is dedicated to my parents.
# Table of Contents

Chapter 1: Introduction ............................................................... 1

Chapter 2: Review of the literature ............................................... 9

Introduction ................................................................................ 9

Theoretical foundations ............................................................... 9

The Interaction Approach in SLA ................................................. 10

Historical overview ................................................................. 11

Components of the Interaction Approach ...................................... 16

Input ....................................................................................... 16

Output ..................................................................................... 17

Corrective feedback ................................................................. 18

Summary .................................................................................. 23

Attention and Awareness in SLA ................................................ 24

Models of attention in SLA ........................................................ 26

Awareness and learning ............................................................. 28

Operationalizing and measuring awareness .................................. 29

Verbal protocols ........................................................................ 30

Using verbal protocols in classroom research ................................. 32

Summary .................................................................................. 38

Effectiveness of corrective feedback in language learning ............... 40

Linguistic target ....................................................................... 41
Type of feedback ................................................................. 44
Direction of the feedback .................................................. 45
Learner proficiency .............................................................. 45
Chapter 3: Design and methods ........................................... 48
Introduction ................................................................. 48
Summary of research ........................................................ 48
Research questions and hypotheses .................................... 49
Operationalizations .......................................................... 52
Corrective feedback .......................................................... 53
Proficiency ................................................................. 53
Awareness ................................................................. 56
Design and procedure ...................................................... 57
Context ................................................................. 59
Observations and participant selection ............................... 61
Participants ................................................................. 62
Informed consent procedure ............................................. 64
Background questionnaire .................................................. 65
Classroom interaction activities ......................................... 65
Data collection procedure .................................................. 68
Data preparation for quantitative and qualitative coding ......... 71
Quantitative analysis of the data .......................................... 73
Theoretical and methodological implications .................................................. 144
Pedagogical implications ................................................................................. 144
Limitations ........................................................................................................ 146
Suggestions for future research ................................................................. 149
Conclusions ....................................................................................................... 149
Appendix A: Transcription and glosses ......................................................... 152
Appendix B: Consent form .............................................................................. 154
Appendix C: Learner questionnaire ............................................................... 157
Appendix D: Instructions for stimulated recall ............................................. 159
References ....................................................................................................... 160
LIST OF FIGURES

Figure 1. Study variables................................................................. 52
Figure 2. Procedure................................................................. 58
Figure 3. Coding categories for quantitative analysis...................... 72
Figure 4. Qualitative analysis – organization of themes.................... 90
Figure 5. Distribution of feedback by type ...................................... 95
Figure 6. Distribution of feedback by target.................................... 97
Figure 7. Beginning learners’ awareness of correction by type of feedback... 104
Figure 8. Advanced learners’ awareness of correction by type of feedback... 106
Figure 9. Advanced learners’ awareness of correction by linguistic target of feedback.. 107
Figure 10. Beginning learners’ awareness of correction by target of feedback........... 108
Figure 11. Beginning and advanced learners’ awareness of correction by direction of feedback................................................................. 110
Figure 12. Beginning and advanced learners’ awareness of target by type of feedback... 112
Figure 13. Beginning and advanced learners’ awareness of target by linguistic target of feedback................................................................. 114
Figure 14. Beginning and advanced learners’ awareness of target by direction of feedback.................................................................................. 115
Figure 15. Global themes identified in the thematic analysis ................. 117
Figure 16. Thematic network “Learner”............................................ 118
Figure 17. Thematic network “Feedback Episode”............................. 122
Figure 18. Thematic network “Linguistic information”. 124

Figure 19. Thematic network “Message”. 128

Figure 20. Thematic network “External distractors”. 129
LIST OF TABLES

Table 1. Veridicality issues in the use of stimulated recall protocols .................. 39
Table 2. Class information .................................................................................. 66
Table 3. Reliability of Coding ............................................................................. 89
Table 4. Types of feedback in recalled corrective feedback episodes ................. 94
Table 5. Linguistic targets of recalled feedback episodes .................................. 96
Table 6. Awareness of corrective intent by learner’s proficiency level ............... 98
Table 7. Awareness of linguistic target by learner’s proficiency level ............... 99
Table 8. Awareness of corrective intent by learner’s proficiency level ............... 101
Table 9. Awareness of linguistic target by learner’s proficiency level ............... 101
Table 10. Beginning and advanced learners’ awareness of correction by type of feedback ........................................................................................................... 105
Table 11. Beginning and advanced learners’ awareness of correction by linguistic target of feedback ........................................................................................................... 108
Table 12. Beginning and advanced learners’ awareness of correction by direction of feedback ........................................................................................................... 109
Table 13. Beginning and advanced learners’ awareness of target by type of feedback. ........................................................................................................... 111
Table 14. Beginning and advanced learners’ awareness of target by linguistic target of feedback ........................................................................................................... 113
Table 15. Beginning and advanced learners’ awareness of target by direction of feedback ........................................................................................................... 114
CHAPTER 1: INTRODUCTION

Corrective feedback is a central issue for teachers and learners in formal instructional settings. On the one hand, feedback in such settings is frequent and widespread (Ortega, 2009) and both groups expect it and consider it a hallmark of good instruction (Bell, 2005; Brown, 2009; Kuntz, 1997b; 2000; Schulz, 1996; 2001). On the other hand, no consensus exists between and among teachers, learners, and language acquisition researchers on the best ways for such feedback to be provided (e.g., Bell, 2005; Schulz, 1996; 2001). For many beginning and experienced teachers the search for efficient, effective, and supportive approaches to addressing their learners’ non-targetlike production, particularly in oral communication, raises important questions about the nature of oral corrective feedback and its application in the language classroom: Is feedback useful? Is it necessary for language development? Do learners understand the feedback they receive – and to what extent? How should feedback be timed, shaped and delivered in order to achieve the best results?

These questions have been the focus of a continuously expanding body of empirical research within the framework of the Interaction Approach (Gass, 1997; Gass & Mackey, 2006; Long, 1996; 2006; Pica, 1994) over the past three decades. Researchers have examined the efficacy of teacher-provided corrective feedback during oral interaction. Findings from developmental studies, conducted both in laboratory conditions (e.g., Adams, 2007; Ayoun, 2001; Egi, 2007; Han, 2002; Ishida, 2004; Iwashita, 2003; Seon Jeon, 2007; Leeman, 2003; Loewen & Nabei, 2007; Long, Inagaki & Ortega, 1998; Mackey, 1999; Mackey & Philp, 1998; Mackey, Philp, Egi, Fujii & Tatsumi, 2002; Mackey & Silver, 2005; McDonough, 2005;
McDonough, 2007; McDonough & Mackey, 2006; Philp, 2003; Van den Branden, 1997; etc.) and in classroom contexts (e.g., Ammar, 2008; Ammar & Spada, 2006; Carroll & Swain, 1993; Doughty & Varela, 1998; Ellis, R., 2007; Ellis, R. & He, 1999; Ellis, R., Loewen & Erlam, 2006; Ellis, R., Tanaka & Yamazaki, 1994; Havranek & Cesnik, 2001; Havranek, 2002; Loewen, 2005; Loewen & Philp, 2006; Lyster, 2004; Mackey, 2006; Muranoi, 2000; Sheen, 2007; etc.) have demonstrated that, in general, feedback plays a facilitative role in language development.

Moreover, empirical research on the efficacy of corrective feedback has already reached the critical mass that allows for research synthesis in the form of meta-analyses to be conducted (Norris & Ortega, 2007). Findings from several such meta-analyses have shown that in comparisons between feedback conditions and no feedback conditions (ignoring errors) the learners who received feedback performed better on post-tests than learners who received no corrections, thus providing further support for the claim that corrective feedback is beneficial for language development (Li, 2010; Mackey & Goo, 2007; Russel & Spada, 2006).

Despite the growing empirical base for the understanding of the role of corrective feedback in second language acquisition, much remains unknown about the factors that moderate the positive effect of corrective feedback on language development. More research is needed that investigates the effect of learner variables, specifically age, gender, and proficiency as well as context-dependent variables, for instance instructional contexts (second language versus foreign language learning, computer-assisted versus traditional delivery of instruction, etc.), target language, particularly languages other than English, and instructional focus (form-focused versus meaning-based instruction) among others (Li, 2010).
One important line of research into the factors that moderate the success of oral corrective feedback is the exploration of the learners’ mental processes of engagement, specifically the learners’ noticing, awareness and comprehension of the feedback. Understanding these mental processes could potentially enable us to manipulate the feedback in order to increase its benefits, i.e. to tailor it to the learners’ capacities, needs, and preferences, in order to maximize its facilitative effect for language learning. Research on these topics has attracted more and more interest in the past decade (see Mackey & Gass, 2006, for overview). Findings so far suggest that the learners’ ability to notice the feedback is moderated by several factors, including characteristics of the learner, characteristics of the instructional context, and characteristics of the feedback. This research, however, is still scarce. There is a particular dearth of studies exploring authentic oral discourse in classroom settings rather than on dyadic interaction in laboratory conditions, as well as of investigations focusing on learners of languages other than English, and learners of varied proficiency. The goal of the present study is to address this gap by employing a mixed-methods exploratory design in Arabic foreign language classrooms at the beginning and advanced levels, in order to shed more light on learners’ mental processes in feedback processing in relation to characteristics of the feedback and characteristics of the learners.

Corrective feedback has been argued to be particularly effective in foreign language settings, where learners tend to focus on and rely on grammatical instruction and the teacher’s corrections more than their counterparts in second language learning environments (Li, 2010; Loewen, Li, Fei, Thompson, Nakatsukasa, Ahn & Chen, 2009; Lyster & Mori, 2006). One reason for the increased importance of feedback in a foreign language learning situation is that the opportunities for extensive spoken interaction outside of the classroom are often limited. These considerations are particularly relevant in the case of Arabic language learning, both in
foreign and in some second language instructional environments. The focus of Arabic language instruction in US institutions and elsewhere is frequently the formal variety, Modern Standard Arabic (MSA) (Angelescu, 2006; Hee-man & El-Khazindar, 2006; Nahla, 2006). MSA is an instructed variety even for native speakers of Arabic, who habitually use their community’s dialects instead for everyday informal communication. Therefore a learner of Arabic – that is, of MSA, would have few opportunities for meaningful communicative interaction in their target language outside the classroom, even if they do come into contact with native speakers of Arabic (Wilmsen, 2006). This leads to an increased emphasis on the potential benefits of the corrective feedback learners receive within the classroom.

As important as its role in the learning of Arabic may be, corrective feedback in the Arabic language classroom has received little attention so far. The few studies conducted in recent years have addressed Arabic learner beliefs, attitudes, and perceptions in relation to error correction (Kuntz, 1997a; 1997b; Loewen, et al., 2009; Mackey, Al-Khalil, Atanassova, Hama, Logan-Terry & Nakatsukasa, 2007). These studies’ contribution to the understanding of the role of feedback in the learning of Arabic notwithstanding, they have been limited in scope (Mackey et al., 2007), have addressed error correction only marginally as part of a larger investigation (Kuntz, 1997a; 1997b) or have included Arabic learners as part of a larger multi-lingual group of participants which makes it difficult to draw conclusions relevant specifically to these learners (Loewen et al., 2009). Further studies, involving a larger number of participants and focusing exclusively on feedback in the Arabic language classroom, remain necessary.

Over the past decade, in response to a rapidly changing political and socio-economic climate, as well as the emergence of new US national security concerns after 11 Sept 2001, Arabic has become a language of high importance and much attention has been focused on
expanding the existing instructional base and maximizing the results from Arabic language learning. According to the results from a comprehensive survey of foreign language enrollments, conducted by the Modern Language Association (MLA, 2010), Arabic – a less commonly taught language in the past – is now the seventh most commonly taught language in the US, marking steady annual increases in enrollments for the past ten years.

Yet despite the rapid increase in the number of learners and teachers engaged in classroom Arabic language instruction over the past decade, empirical research on the learning and teaching of Arabic has not kept up with the pace of enrollments (al-Batal, 1995; England, 2006; Ryding, 2006). Since Arabic language classes have witnessed soaring enrollments but also high attrition, research on the learning processes in the Arabic classroom becomes increasingly urgent. To maximize the learning potential of the growing classroom enrollments, it is important to focus research efforts on understanding how learners of Arabic learn, and what can be done to help them learn better, faster, and more efficiently.

Awareness of the need for more empirical research into the teaching and learning of Arabic has increased in recent years, a tendency that has been reflected in a growing number of dedicated thematic fora, as well as increased published research. In 2008, the American Council for the Teaching of Foreign Languages (ACTFL) sponsored the creation of a Special Interest Group (SIG) for Arabic, which has been tasked with organizing language-specific research panels as part of the ACTFL annual meeting. In the same year, two conferences – the National Arabic Language Conference (June 1-15, 2008, DePaul University, Chicago, Illinois) and the Action Research Meeting Conference on K-12 Arabic Learning and Teaching (15 Nov, 2008, Michigan State University, East Lansing, Michigan) – provided a forum and opportunities for
collaborations to researchers conducting theoretical and empirical work on the acquisition of Arabic as a second, foreign, or heritage language.

There has also been a steady, albeit small, increase in the quantity of published research focusing on the acquisition of Arabic, particularly studies focusing on language universals and cognitive processing (AlHawary, 1999; 2009; Mansouri, 2005; Nielsen, 1997), the role of affect and learner strategies in listening (ElKhafaifi, 2005; 2007), reading in Arabic (Mughazi, 2005), the development of highly advanced capabilities (Samimy, 2008), classroom interaction (Mackey et al. 2007), learner needs, beliefs and demographics (Belnap, 2006; El Essawy, 2006; Kuntz, 1997a; 1997b; 2000), and dialect comprehension (Trentman, 2011).

Despite these optimistic developments, the existing body of published and accessible empirical research remains far from reaching the critical mass that is necessary for research syntheses in the form of meta-analyses to start taking place (Ortega & Norris, 2007). It is this sort of synthesis of empirical findings across settings, methodologies, and populations that can ultimately provide a sound theoretical motivation for the selection and promotion of learning contexts, teaching and assessment approaches, and instructional standards.

The goal of this study is, therefore, twofold. First, it will add to the body of empirical research on the factors constraining the efficacy of corrective feedback, by investigating the interplay between beginning and advanced learners’ awareness of corrective feedback during classroom interaction on the one hand, and the linguistic, typological, and contextual characteristics of the feedback (type, linguistic target, and direction of the feedback) on the other.

Second, it will contribute to the ongoing empirical shift in the field of Arabic language instruction by using a mixed-method, multifaceted approach to explore classroom interaction
from the point of view of learners and teachers, in order to provide new insights into the process of learning Arabic in a formal foreign language classroom setting.

This dissertation consists of five chapters, which will be described below.

Chapter 1, Introduction, provided an overview of the research by giving a brief theoretical background to explain the motivation of the study, formulating the research problem, describing the methodology employed to answer the research questions, and addressing the significance of the study.

Chapter 2, Review of the Literature, introduces the theoretical framework for the study and surveys the existing empirical research on corrective feedback and awareness.

Chapter 3, Design and Methods, describes the methodological framework of the study, the participants, materials, procedures for data collection, and the qualitative and quantitative methods employed in the analysis of the data.

In Chapter 4, Results, the findings from the quantitative and qualitative analysis of the data are presented. Descriptive and inferential statistics are reported for each quantitative research question and findings from the thematic analysis of the narrative data are used to provide richer insights into learner understanding and motivation.

In the final Chapter 5, Discussion and Conclusions, the results of the quantitative analysis, supplemented with insights from the qualitative exploration of the data, are discussed in relation to relevant theoretical and empirical research. The findings are summarized and their theoretical, methodological, and pedagogical implications discussed. The chapter concludes with
an outline of the limitations of the study, followed by suggestions for considerations to be made in future research.
CHAPTER 2: REVIEW OF THE LITERATURE

Introduction

This chapter presents the theoretical framework for the study and introduces the main theoretical constructs addressed therein (feedback and awareness). It is organized as follows: The first section will present the theoretical framework of the Interaction Approach, its historical foundations, current framework, and empirical base. The second section will introduce the constructs of attention and awareness, and discuss their place in the learning process. The following section will deal with the current methodological issues related to defining and measuring awareness in an interactive learning situation, and present a theoretical justification for the awareness measure selected for this study. In section four, the role of awareness in the processing of corrective feedback as one of the elements of the Interaction framework will be discussed in relation to empirical findings from laboratory and classroom studies on language learning. The factors, constraining the learners’ ability to notice feedback will be identified, and the gaps in the literature which motivated the current study will be summarized.

Theoretical foundations

This study was conducted within the theoretical framework of the Interaction Approach from a predominantly cognitive-interactionist perspective. Cognitive interactionism is associated with Piagetian developmental psychology (e.g., Piaget, 1974), and refers to the position that learning is the result of the interaction between learner- internal cognitive factors and learner-
external social or environmental factors (Ortega, 2009). These factors are taken to be separate, rather than overlapping, and their interaction is presupposed to bring about learning in the form of learner-internal cognitive changes.

The cognitive-interactionist theoretical framework stands distinctly apart from the more recent socio-cultural perspective that was introduced in SLA in the late 90s with the application of social theories to the explanation of the language learning process (e.g., Vygotskian sociocultural theory, Systemic Functional Linguistics, language socialization theory and identity theory). This socio-cultural perspective views learning as a fundamentally social phenomenon, and posits that the exploration of individual, socially-shaped perspectives ultimately carries more explanatory power for the complex problem of learning than the search for universal principles and cognitive-individual constructs. Despite the seemingly irreconcilable differences between these two theoretical stances, in recent years they have increasingly been brought together to provide broader explanations for learning phenomena under the most recently articulated framework of the Interaction Approach. An example of this approach are the edited volumes by Han (2008) and Mackey & Polio (2009). This is the approach adopted in the current study as well: While the overall concept, design, and measurements are strictly cognitive-interactionist, insights from socio-cultural perspectives have also been applied in the interpretation of the study’s findings.

The Interaction Approach in SLA

The main premise of the Interaction Approach (Gass, 1997, 2003; Gass & Mackey, 2006; Long, 1996, 2006; Pica, 1994) is that it is through interaction with interlocutors that language
learners have opportunities to process linguistic input in the form of positive and negative evidence, and to notice differences between their own formulations of linguistic structures in the target language, and those of their conversational partners. Corrective feedback, received as part of this interaction, particularly when noticed and recognized as a gap between the learner’s production and the target language, both modifies the linguistic input learners receive, and pushes them to modify their own output (Gass & Mackey, 2006).

**Historical overview**

The Interactionist perspective stems from Hatch’s seminal work in the 1970s and the early 1980s. Until that time, conversation in the target language was primarily viewed as an opportunity to practice linguistic structures, already learned in the classroom. In a series of studies (Hatch, 1978; 1983; Hatch, Deck & Wagner-Gough, 1979), Hatch demonstrated that conversation may, in fact, play the leading role in the learning process, with L2 syntax actually evolving out of conversation, as opposed to being the product of explicit learning which conversation simply refines. Long (1983) investigated the patterns of conversational interaction between native speakers (NS), and between native speakers and non-native speakers (NNS). He found higher frequencies of interactional modifications, particularly clarification requests, comprehension checks, and confirmations, in NS-NNS interactions, which led him to conclude that NS-NS and NS-NNS interactions are fundamentally different in the amount of modification to the linguistic input that they provide.

At the same time as this early work, Krashen’s Input Hypothesis (1985) was formulated. The Input hypothesis claims that comprehensible input is a necessary and sufficient condition for language acquisition to take place, provided that the affective filter (a combination of affective
variables such as anxiety) is lowered enough for the comprehensible input to be accessed and processed by the learner. Long (1985) attempted to establish an empirical link between the interactional modifications he had found in NS-NNS conversation, and language acquisition, by testing whether modified input (i.e., input that is simplified or otherwise tailored in response to the learner’s conversational moves), would promote comprehension – and thus, indirectly, impact learning. His results showed higher comprehension scores for learners who received modified input, than for learners in the control condition. Under the premise of Krashen’s Input hypothesis, Long claimed that interactional modifications make input more comprehensible to learners, thus promoting second language acquisition.

The first attempt to test directly the link between interaction and second language development was Sato’s (1986) longitudinal observational study of two Vietnamese boys, learning English in a naturalistic setting. Her work brought to the fore a more complex picture of the relationship between interaction and learning, than had been hypothesized before. For example, in tracking the morphosyntactic development of her participants over a 10-month period, she noted that they showed little improvement in the use of the past tense inflection. Further analysis clarified that the two boys chose to circumvent the troublesome forms, by relying heavily on contextual cues, as well as lexical and adverbial means for pinpointing past time actions. This strategy rendered the recognition and productive use of past tense morphology non-essential for their ability to interact in a conversation. On the basis of such findings, Sato concluded that conversational interaction selectively facilitates the learning of some linguistic forms but not others, and that, while interaction plays a role in language development, it alone does not explain the ultimate achievement of native-like proficiency in a second language.
During the same period, Swain (1985) introduced a new variable in the already complex picture of interaction-based language learning. Her work on the language development of French L1 school-children in immersion ESL classrooms in Canada showed that comprehensible input alone is insufficient to promote the development of L2 morphosyntax. In her Output Hypothesis (1985, 1993, 1995), she argues that an active communicative role for the language learner is essential to their development, i.e. “through producing language, either spoken or written, language learning may occur” (Swain, 1993, p. 159). Swain asserts that output is essential for pushing the learner to switch from semantic processing to syntactic processing. While semantic processing alone allows the learners to ignore syntactic features of little value for comprehension, pushing them to produce output forces them to attend to those less salient features, as they attempt to formulate precisely their intended meaning. The process of formulating output further forces the learners to “notice what they do not know, or know only partially” (Swain, 1995, p.129). Being aware of the gap in their interlanguage may in turn force the learners to attend to linguistic input more, in order to fill the gap (Swain, 1993). This claims has been echoed by other researchers (e.g., Gass, 1997; White, 1987) who argue for the need for incomprehensible input, rather than comprehensible input. They contend that it is the failure to comprehend that eventually sensitizes the learner to the gap in their knowledge and to realization of their problem areas (White, 1987) and that it is this realization that triggers the processes of interlanguage restructuring.

The above factors, identified as driving forces in interaction-based learning, were synthesized in Long’s updated Interaction Hypothesis (1996): that “negotiation for meaning, especially negotiation work that triggers interactional adjustments by the NS or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities,
particularly selective attention, and output in productive ways” (pp. 451-452). In this view of second language acquisition, negotiation of meaning plays a major role. Negotiation of meaning is defined as “the modification and restructuring of interaction that occurs when learners and their interlocutors anticipate, perceive, or experience difficulties in message comprehensibility” (Pica, 1994, p. 494). During negotiation of communication breakdowns, learners are pushed to make their speech more target-like, facilitating the acquisition of L2 linguistic features that may otherwise be ignored (Swain, 1985, 1993, 1995). Attempts to produce more target-like output may highlight a gap in the learner’s interlanguage, and consequently lead to internal linguistic restructuring (Gass, 1997). Finally, the processes of input and output modification expose the learners to higher frequencies of redundant linguistic input, making the noticing of non-salient features, and their subsequent learning, more likely (Long, 1996). To summarize, participating in conversational interaction allows learners to focus on the linguistic form of the input, “process the input they receive, indicate what modifications are needed to make that input comprehensible, formulate output, receive cues about the targetlikeness of their output, and use cognitive abilities like selective attention in order to connect the input they receive to the output they produce” (Adams, 2004).

A large volume of empirical research has been conducted within the framework of the Interactionist approach in SLA. The early studies were primarily observational in nature and sought to establish descriptive accounts and taxonomies for interactional modifications (e.g., Gass & Varonis, 1985; Long, 1981, 1983), as well as the factors affecting the type and amount of negotiation, including task type (e.g., Doughty & Pica, 1986; Foster & Skehan, 1996), task familiarity (e.g., Plough & Gass, 1993), interlocutor gender (e.g., Gass & Varonis, 1986; Pica,
Holliday, Lewis, Berducci & Newman, 1991), and the type of interlocutor, e.g. NS versus NNS (e.g., Gass & Varonis, 1985; Pica, Lincoln-Porter, Paninos & Linnell, 1996).

Starting in the mid-1990s, the research focus shifted to empirical investigations of the link between interaction and L2 learning (Mackey & Gass, 2006). Positive effects for interaction on comprehension and language development were demonstrated in the majority of the studies (e.g., Ellis, Tanaka & Yamazaki, 1994; Gass & Varonis, 1994; Iwashita, 2003; Mackey, 1999, Mackey & Oliver, 2002; Pica, Young & Doughty, 1987; Polio & Gass, 1998, etc.). One notable exception to the line of research that established an empirical link between interaction and learning, was Loschky (1994), who found no positive effects of interaction on the learning of Japanese locative structures and vocabulary items. Loschky’s study highlighted the need to explore further the factors that constrain the positive effect of interaction on language development, including the context in which the interaction took place, the linguistic structures at the core of the interaction, learner-specific factors, etc.

Currently, the Interactionist Approach is firmly established as a productive theoretical account for SLA (Gass & Mackey, 2006; Mackey & Gass, 2006). Research today has moved “beyond the question of whether interaction plays a role in development to asking how it facilitates development” (Mackey & Gass, 2006:171, emphasis mine). New research techniques and methods, such as the expanding use of computer-based treatments and data-collection methods, as well as increasingly sophisticated techniques for the study of learner-internal cognitive and perceptual processes, ranging from recall and think-aloud verbal protocols to emergent instruments in cognitive neuroscience, such as eye-tracking, electroencephalography (EEG), and functional magnetic resonance imaging (fMRI), have opened new paths for empirical inquiry (Mackey, 2006).
Components of the Interaction Approach

In sum, the Interaction Approach, which has been developing as a theoretical account of SLA since the late 70s, is now an established theoretical and empirical research paradigm. The framework proposed by this Approach as explanation for the processes taking place in SLA, can be described as a constant interplay between the linguistic input that the learners are exposed to, the output formulated by the learner, the feedback regarding the targetlikeness of the learner’s output, and, finally interactional modifications to both the input and the output, triggered by conversational breakdowns, which eventually lead to restructuring of the learners’ interlanguage. In the following pages, each of these components of the Interaction Approach will be presented and discussed.

Input

Input refers to the linguistic forms to which the learners are exposed. Most theories of second language learning acknowledge the significance of input as a basic component in the learning process. Input is the “raw data” which serves as linguistic evidence to show learners how to formulate linguistic output in the target language system.

Although the significance of input is widely recognized by most accounts of SLA, it does not carry the same weight in all of them. For example, the Universal Grammar framework limits the role of input to the “trigger that interacts with an innate system and/or the native language to promote language learning” (Gass & Mackey, 2006, p. 5). Frequency-based approaches to SLA (e.g., Ellis, N., 2002) assign input a central role in the learning process, which they view as a combination of exposure to input and the use of input in communication.
Because input has a significant role to play in language learning, SLA researchers have sought to define how the input needs of L2 learners can be met – i.e., what kinds of input and in what form are most facilitative for language learning. Research so far has addressed the differing developmental outcomes from the exposure to simplified input (language that has been de-complexified so as to be more comprehensible), baseline input (unmodified language), or interactionally modified input (input that is modified in the process of conversational interaction with an interlocutor, as a result of indicated communicative breakdown) (Gass & Mackey, 2006; Mackey & Abbuhl, 2005).

A number of studies have sought to compare the relative effectiveness of input simplification and interactionally modified input on second language comprehension and learning (e.g., Ellis & he, 1999; Ellis, Tanaka & Yamazaki, 1994; Gass & Varonis, 1994; Loschky, 1994; Mackey, 1999; Pica, 1992; Pica, Young & Doughty, 1987). The results of these studies indicated that interactionally modified input may be more effective than simplified input for language development. Proponents of the Interaction Hypothesis (Long, 1996) suggest that this may be due to the fact that during interactional modifications learners have the opportunity to negotiate the type of input that best fits their particular developmental stage. Simplified input, on the other hand, is modified beforehand, and does not show the same sensitivity to “individual learners’ particular weaknesses, strengths, and real-time communicative needs in relation to the target language” (Mackey & Abbuhl, 2005, p. 210).

Output

Output refers to the language that learners produce. Some researchers make a distinction between comprehensible output (utterances that are modified so as to be understandable to the
learner’s interlocutor) and modified output (utterances that have been changed to be more-targetlike in response to an interlocutor’s signal of communicative breakdown). The latter type is the typical focus of interaction research (e.g., Ellis & He, 1999; McDonough, 2005; McDonough & Mackey, 2008; Muranoi, 2000; Shehadeh, 2002; Swain, 1985, 1995, 2005). This sort of output has been attributed a facilitative role in language learning since it pushes the learners to reflect on their original production and notice gaps between it and the target language. Researchers have argued that output gives the learner opportunities to receive feedback, which in turn stimulates them to produce more accurate, linguistically complex, and comprehensible forms, as well as to formulate hypotheses about the target language, and modify them according to the conversational response they receive (e.g., Swain, 1993; 2005; Swain & Lapkin, 1995; Gass, 1988, 1997; Long, 1996; Pica, 1994).

Output is hypothesized to benefit language learners in four important ways: it promotes fluency; it draws learners’ attention to linguistic problems; it encourages a gear-shift from semantic processing of L2 input (i.e., focus on the meaning) to syntactic processing (i.e., focusing on the linguistic structures needed to express their intended meaning, and, finally, it allows learners to form and test hypotheses about the structure of the target language and what is – or is not – possible within the constraints of the target language’s linguistic system (Mackey & Abbuhl, 2005; Ellis, R., 2009).

**Corrective feedback**

Corrective feedback refers to information that learners receive in response to their communicative efforts – typically, regarding non-targetlikeness of their output, and/or communicative breakdowns associated with such non-targetlikeness. Feedback is believed to be
an important part of the language learning process, with researchers arguing that certain types of linguistic structures are not directly acquirable from linguistic input alone (e.g. White, 1991) for reasons including inter-language influence, which obfuscates certain linguistic forms that are both unmarked in the learner’s native language, and non-salient in the target language input. In the past decade, a number of research studies have addressed the relationship between oral feedback, particularly implicit negative feedback (recasts), and language learning (e.g., Braidi, 2002; Han, 2001, 2002; Iwashita, 2003; Leeman, 2003; Long, Inagaki & Ortega, 1998; Lyster, 1998a, 1998b; Lyster & Ranta, 1997; Mackey, 1999; Mackey, Gass & McDonough, 2000; Mackey & Philp, 1998; Mackey, Oliver & Leeman, 2003; Oliver, 1995, 2000; Philp, 2003, etc.).

Explicit feedback

In second language classrooms, language teachers use a wide range of corrective feedback to help learners identify problems in their non-targetlike utterances. Example 1 illustrates explicit feedback (all examples come from data collected for the current study; for transliteration conventions, abbreviations and symbols, see Appendix A).

Example 1. Explicit Feedback

T: \textit{DHanna, DHanantu – wa hum?}  
(He-thought, I-thought, and they?)

He thought, I thought, and they?

L: \textit{yaDHunnuun}  
(They-think)

They think

T: Past tense, past tense, \textit{maaDii} (Explicit feedback)
L: **DHannuu**

(They-thought)

They thought

In this example, the teacher began by eliciting a verb form. When the form was supplied in an incorrect tense, she provided feedback explicitly by asking the learner to use the past tense.

Explicit feedback was the first type of feedback to receive attention in empirical research, and its role in language development remains controversial. Studies on explicit feedback addressed the question of whether this type of feedback affected performance (e.g., Carroll & Swain, 1993; Carroll, Swain & Roberge, 1992; DeKeyser, 1993; Herron & Tomasello, 1988; Lightbown & Spada, 1990, 1993; Tomasello & Herron, 1989). For example, one treatment-posttest-delayed posttest laboratory study, Carroll et al. (1992) examined the effect of explicit feedback on the learning of French derivational morphology. The results showed that explicit feedback helped learners of French as a second language to memorize words, but not to learn the derivational rules that guided them. However, in laboratory studies with similar design, Carroll and her colleagues (Carroll & Swain, 1991, 1993) found that explicit feedback was, in fact, more beneficial for Spanish ESL learners acquiring the dative alteration rule than implicit feedback or no feedback at all. Carroll and Swain hypothesized that the reason for the relative effectiveness of explicit feedback in their study was that it helped learners pinpoint the exact location of the error, rather than to guess its nature and location. Similar results were reported by Kubota (1994) and DeKeyser (1993). Overall, empirical findings from research on the effectiveness of explicit feedback remain inconclusive, mostly due to the theoretical and pedagogical shift to implicit forms of instruction in the 1990s. Explicit feedback and explicit instruction have only recently
resumed their place as research foci, largely due to a renewed interest in the mechanism and processes involved in the development of native-like language capabilities.

*Implicit feedback*

Implicit feedback does not overtly mark the learner’s production as non-targetlike. In meaning-oriented language classrooms, teachers have been found to be more likely to use implicit than explicit feedback (e.g., Ellis, Basturkmen, & Loewen, 2001; Lyster & Ranta, 1997, Seedhouse, 1997). Implicit feedback may take the form of a recast, or a more targetlike version of what a learner has just said. Recasts can be declarative, as in Example 2, in which the teacher provides the target-like form with declarative intonation.

Example 2. Declarative Recast.

L: *li-ʾannahu wajada waDHiifa *munaasib* (Morphological error-agreement)  
   (because-he he-found job *suitable-masc)  
   because he found a *suitable job

T: *munaasiba*  (Recast)  
   (suitable-fem)  
   suitable

In addition to the declarative form, a recast can also be provided in interrogative form, that is, given with interrogative intonation, as in Example 3 below.

Example 3. Interrogative Recast.

L: *huwa mutaxaSSiS fii ... fi-l.... *fi-l-fiziika?* (Lexical error)  
   (he-specializes in in-the in-the-*fizika?)  
   he specializes in *fizika?
T: *fiiziyaa’?* (Interrogative recast)

Physics?

In this example, the learner pronounced the lexical item in a non-targetlike way. The teacher provided the correct form with interrogative intonation.

Another common form of implicit feedback is negotiation, shown in Example 4.

Example 4. Negotiation

L: *saghiir ʾuxt ʿindii .. ʿindii *saghiir ʾuxt ʿinduki?

(*little-masc sister I-have I-have *little-masc sister you have?)

I have… younger sister…I have…do you have?

T: hal ʿindii…? (Clarification question – communicative breakdown)

do I have…?

L: ʿinduki *saghiira ʾuxt?

(you-have little *sister little?)

do you have a *sister younger?

T: ʾuxt saghiira (Recast)

younger sister

Negotiations typically occur when a teacher does not understand a learner’s utterance. In some negotiations the teacher may repair communication breakdowns by repeating the learner’s problematic utterance as in Example 4. A teacher may also negotiate by asking clarification questions without repeating any part of a learner’s utterance. In Example 4, the negotiation sequence resulted in the learner changing the non-targetlike formulation *saghiira ʾuxt* to the targetlike syntactic formulation of ʾuxt saghiira meaning “younger sister”.

22
Implicit feedback such as recasts and negotiation, has been the focus of much empirical research over the past two decades. Generally, researchers have focused on two main empirical questions: whether implicit feedback promotes learning, and whether implicit feedback is recognized as such by the learners. The latter question stems from a line of research, based on Schmidt’s noticing hypothesis (Schmidt, 1990; 1995; 2001; 2010), claiming that the usefulness of feedback is directly related to whether or not it is noticed by the learner (e.g. Long, 1996; Roberts, 1995; Mackey et al, 2000). This research will be examined in more detail in the next section.

Summary

Empirical research on the role of interaction in language learning over the past three decades has been expanding both in terms of quantity, and in terms of the scope of research questions and research methodologies. Overall, findings from this research posit a beneficial role for interaction in language learning.

Developmental studies have demonstrated that, in general, interaction, particularly conversational interaction that involves the provision of corrective feedback, plays a facilitative role in language development (see meta-analyses by Li, 2010; Mackey & Goo, 2007; Russel & Spada, 2006).

Attention, awareness, and noticing have been proposed as factors moderating the effectiveness of interaction for language development. Schmidt’s Noticing Hypothesis (Schmidt, 1990; 1995, and elsewhere) assigns noticing – defined as detection and awareness – a central and indispensable role in learning. The next section will deal with the constructs of attention and awareness in SLA and their role in language development.
Attention and Awareness in SLA

Most theoretical accounts for SLA posit some role for attention; however, this construct is, naturally, especially emphasized in interactionist-cognitivist accounts (Bowles & Leow, 2005). Schmidt (2001) goes as far as to claim that “attention appears necessary for understanding nearly every aspect of second and foreign language learning” (Schmidt, 2001, p. 6). This view is supported by a long-standing and empirically supported position in cognitive psychology, according to which little, if any, learning can take place without attention, since attention to stimuli is necessary for long-term memory storage (Carr & Curran, 1994; Nissen & Bullemer, 1987; Posner, 1992; Reber, 1976; 1989). In SLA, the theoretical models related to the role of attention have been largely developed in response to theoretical work in cognitive psychology. Three major theoretical models of attention, accepted in cognitive psychology, have been applied within the field of SLA: the filter model, the effortful attention model, and the unlimited capacity model.

The classical view of human attention, presented in the cognitive psychology literature, is that human attention is a limited-capacity resource and that humans are limited-capacity processors (e.g., Broadbent, 1958; Kahneman, 1973). Since human attention was limited, it was also, by necessity, selective (i.e., not all stimuli received attention). Several theories were proposed regarding the nature and timing of this selectivity. Filter theories of attention proposed that a certain filter (bottleneck) existed in the central processor, that only allowed some of the stimuli to be processed. This filtering happened very early in the process, that is, stimuli were either selected for processing or not, as soon as they occurred – hence the name of this bottleneck
model, “early selection model” (Broadbent, 1958). Once past the bottleneck and admitted for processing, stimuli were processed serially (i.e., one after the other, rather than simultaneously), so that attention to one piece of information blocked attention to another.

Research on simultaneous processing (particularly on dichotic listening, or shadowing) in the 1960s, however, demonstrated that information in the input stream can be processed in parallel, and allowed the formulation of a new model, “late selection model”, according to which all stimuli were initially accepted and processed parallelly, and selection of input to attend to happened at a later stage (Deutsch & Deutsch, 1963).

In contrast to the filter models, which describe the human processor as a largely passive recipient of information, recent models have posited a more active role for the information processor, in the form of effortful selective attention. Two major models of effortful attention describe the process of information selection as either a single-pool or a multiple-pool resource process. For example, Kahneman (1973) proposes that attention, while limited, can handle two tasks at the same time, provided that task demands are not too high (i.e., the tasks are automatized), and there is sufficient arousal. Wickens (1980), on the other hand, proposes a model of attention in which attention consist of several resource pools, rather a single one. According to his model, the ability to perform two tasks simultaneously depends on whether the necessary recourses are drawn from the same, or from different pools. He claims that performing two tasks that draw on the same pool (requiring serial processing, e.g. talking and writing at the same time) would be more difficult than performing two tasks that draw on different pools of attentional resources (e.g., talking and driving a car). Performing two tasks that draw on the same pool of attentional resources would only be possible if one of them, or both, are automatized.
The theoretical models of attention, proposed in cognitive psychology, have been reflected in three major theoretical positions in SLA. First, Tomlin & Villa (1994) proposed a functionally-based fine-grained framework for the analysis of attention, which describes attention as comprising three components: alertness (readiness to deal with incoming stimuli), orientation (direction of attentional resources to a selected type of stimuli), and detection (cognitive registration of stimuli). Tomlin and Villa argue that detection is the stage that allows further processing of input and subsequent learning to take place. The other two elements may play a supportive role, but neither of them is actually necessary for learning to occur. Detection in this model does not imply awareness, or a conscious registration of stimuli – that is, according to this model, awareness is not a necessary condition for learning. Recently this model has been criticized for being too general to provide a useful account of attention in language learning (e.g. Schmidt, 2001; Simard & Wong, 2001).

In contrast to Tomlin & Villa’s view that awareness is not a necessary condition for learning, Schmidt proposed his Noticing hypothesis (1990, 1993, 1994, 1995, 2001, 2009), according to which without awareness input can be processed only only in short-term memory and therefore cannot reach sufficient processing depth for registration in the long-term memory, i.e., learning, to occur. According to this hypothesis, attention is isomorphic with awareness, and the key to the conscious noticing of input, which is “the necessary and sufficient condition for the conversion of input into intake” (Schmidt, 1993, p.209). In addition to noticing, Schmidt proposes a second level of awareness, awareness at the level of understanding, which enables such complex functions as analysis, cognitive comparisons, and hypothesis testing, which lead to deeper learning in the form of restructuring and systematic changes to the learner’s
interlanguage. Empirical studies in SLA have provided support for the differential effects of different levels of awareness on L2 development (e.g., Bowles, 2003; Leow, 1997, 2001; Rosa & Leow, 2004; Rosa & O’Neill, 1999). The main problem, raised for the Noticing hypothesis as a theoretical account of the role of attention in language learning, is the fact that this model is inherently unfalsifiable – that is, while it is possible to measure awareness and examine its role, the opposite, complete lack of awareness, is impossible to demonstrate empirically, which undermines the theoretical validity of the model.

A third model of attention in SLA, proposed by Robinson (1995, 1996) reconciles the two positions described above, by combining Tomlin & Villa’s (1994) notion of detection (which does not involve awareness) and Schmidt’s (1990, 1993, and elsewhere) notion of noticing or awareness. In the model proposed by Robinson, detection is an early stage in the process, which occurs prior to noticing. Noticing on the other hand is defined as “detection plus rehearsal in short-term memory, prior to encoding in long-term memory” (Robinson, 1995, p. 296). Noticing, in this model, does involve awareness, and is a crucial stage in the learning process.

Drawing on work in cognitive psychology by Neumann (1996), Robinson (2003) has introduced the unlimited capacity model of attention to SLA, claiming that it may better explain attention-related cognitive processes in language learning. However, as Bowles & Leow (2005) note, “this model will need substantial elaboration and empirical support before it can be afforded the status of the limited capacity models (and especially that of the noticing hypothesis, which is backed by substantial empirical evidence in SLA)” (p. 182).
To summarize, three models for attention have been proposed in the field of SLA: Tomlin & Villa’s detection model which postulates that awareness is not necessary for language learning, Schmidt’s noticing hypothesis, which posits an essential role for awareness, defined as conscious noticing, in language learning, and Robinson’s model which reconciles the two previous models, by assigning both detection and noticing a role at differing stages in the learning process, but ultimately claims that awareness is crucial for learning to take place.

Awareness and learning

As can be seen from the above overview of theoretical models for attention, the facilitative role of some level of attention (whether detection, noticing, or awareness) is generally accepted, but the role of awareness remains controversial. This controversy is a reflection of a larger debate in cognitive psychology on the role of awareness in learning in general. It seems that the allocation of attentional resources to specific stimuli (in the case of language learning – to specific linguistic items in the input), is necessary for further processing, but whether learner awareness, expressed as conscious noticing of the linguistic information is required for the information to be processed, is open to debate. Some researchers claim that awareness is not a necessary condition for learning (e.g., Carr & Curran, 1994; Tomlin & Villa, 1994; Velmans, 1991) while others insist that a relationship between learning and awareness exists (Robinson, 1995, Schmidt, 1990; 2001; 2009).

A number of empirical studies have provided empirical support for Schmidt’s noticing hypothesis and specifically for the role of awareness and the differential effects of distinct levels of awareness in language learning. Overall, these studies appear to support a facilitative effect of awareness on language learning. Specifically, findings from such studies indicate that awareness
at the level of noticing and understanding contributes significantly to increased learner ability to process the targeted form or structure (Leow, 1997, 2000, 2001; Rosa & Leow, 2004, Rosa & O’Neill, 1999). Second, awareness at the level of understanding has been demonstrated to lead to significantly more intake when compared to awareness at the level of noticing (Leow, 1997, 2001; Rosa & Leow, 2004; Rosa & O’Neill, 1999). Third, a correlation has been shown between awareness at the level of understanding and learner use of hypothesis testing or rule formation (Leow, 1997, 2000, 2001; Rosa & Leow, 2004; Rosa & O’Neill, 1999).

Operationalizing and measuring awareness

The operationalization and measurement of what constitutes awareness in SLA is a complex methodological issue (see Leow, 2001, for a review). Until recently, most studies on language development followed a traditional pre-test post-test design, with no operationalization or measurement of attention or awareness. Instead, posttest scores were used to hypothesize whether the treatment had led to increased attention or awareness of the targeted linguistic structure in the data.

In recent years, however, a number of specific measures of awareness have increasingly been employed in SLA research on awareness and attention, including off-line questionnaires (Alanen, 1995; Izumi, 1999; Mackey, Philp, Egi, Fujii & Tatsumi, 2002; Robinson, 1995, 1996, 1997a, 1997b), off-line uptake charts (Mackey, McDonough, Fujii & Tatsumi, 2001), free recall of input (Greenslade, Bouden & Sanz, 1999; VanPatten, 1990), learning diaries (Altman, 1997, Grabe & Stoller, 1997; McDonough, 2008; Schmidt & Frota, 1986; Warden, Lapkin, Swain & Hart, 1995). In addition to those measures, a line of research has used verbal protocols to measure attention and awareness during exposure to linguistic input.
The measures of awareness, employed in SLA today, all stem from Allport’s (1988) three criteria for establishing awareness: of action (the subject demonstrates awareness by exhibiting some behavioral or cognitive change as a result to being exposed to a stimulus); of memory (the subject is able to remember the events or activity later); and of confidence (the subject must be able to indicate correctly and confidently the stimulus that has been presented). (Allport, 1988: 166-171). A number of SLA studies have focused on behavioral or cognitive changes, in the form of uptake or repair in response to interactional feedback, as evidence for awareness. While these measures fulfill Allport’s first requirement for establishing awareness, overreliance on such measurements can challenge the internal validity of a study, since the overt expression of behavioral or cognitive changes in a second language may be severely constrained by other intervening variables like learner proficiency and general linguistic ability. In simplified terms, a learner’s failure to repair their initial non-targetlike utterance after receiving a correction may be an indication that the learner did not notice the correction, but may equally likely be the result of the learner’s linguistic abilities falling short of modifying their output to the necessary extent – or at all. In an effort to assess more directly learner-internal cognitive processes during interaction, researcher have turned to measures like introspective verbal protocols.

**Verbal protocols**

Verbal protocols, simply defined, require the subject to verbalize their thought processes either while carrying a task or a short time after completing a task. Thus verbal reports can be categorized as concurrent (at the same time as the task) or retrospective (some time after), based on the timing of the collection of reports, relative to the task (Ericsson & Simon, 1984, 1993). Two assumptions underlie verbal reports: that it is possible for internal processes to be observed
in the same way as external real-world events can be observed, and that subjects can verbalize the internal thought processes to which they have access.

Verbal reports have been used since the early 1980s in first language research and other fields. A number of studies have used think-aloud protocols to probe students’ reading strategies (e.g., Cohen, 1986, 1987; Earthman, 1992; Folger, 2001; Gordon, 1990) and some have used verbalization as a technique to compare reading strategies (e.g., Fehrenbach, 1991; McGuire & Yewchuk, 1996). In first-language writing research, verbal protocols have been widely used to compare the cognitive processes involved in writing different kinds of texts (Durst, 1987) and investigating the process of revision and editing (e.g., Breetvelt, 1994; Zellermayer & Cohen, 1996). In recent years, verbal protocols, specifically think-alouds, have been applied to the study of language testing.

In addition, introspective measures such as verbal reports have been used extensively as a data elicitation technique in L2 research, including second language reading and writing (e.g., Cavalcanti & Cohen, 1990; Cohen, 1987; Cohen & Cavalcanti, 1987; Hosenfeld, 1976, 1977, 1979, 1984), comparisons between L1 and L2 strategies (e.g., Chamot and el-Dinary, 1999; Davis & Bistodeau, 1993; Nevo, 1989; Yamashita, 2002), L2 test-taking strategies (e.g., Cohen, 2000; Norris, 1992; Warren, 1996), translation (e.g., Enkvist, 1995; Faerch & Kasper, 1986; Jaaskelainen, 2000; Kern, 1994), interlanguage pragmatics (e.g., Cohen and Hoselfeld, 1981; Kasper and Blum-Kulka, 1993), and oral interaction research (e.g., Egi, 2004a, 2004b, 2007, 2008; Mackey, AlKhalil, Atanassova, Hama, Logan-Terry & Nakatsukasa, 2007; Mackey, Gass & McDonough, 2000; Nabe & Swain, 2004; Philp, 2003, Roberts, 1995), to L2 attention and awareness studies (e.g., Leow, 1997, 1998, 1999, 2000, 2001, Rosa & Leow, 2004a, 2004b; Rosa & O’Neill, 1999).
As can be judged from the large number of studies in which verbal reports have been employed as an instrument for measuring awareness, this type of data collection technique is well-established in SLA research.

*Using verbal protocols in classroom research: Issues and solutions*

The first published SLA study which employed verbal reports was Alanen (1995). She used concurrent reports (think-alouds) in combination with off-line measures (questionnaires) to explore the role of visual input enhancement and noticing in the acquisition of structural language elements. Since then, verbal reports have been steadily gaining popularity as a data elicitation technique. Over the past two decades, empirical studies on the role of attention in SLA have employed either concurrent or retrospective verbal reports as measurements of awareness or noticing (Egi, 2004a, 2004b, 2007, 2008; Leow, 1997, 1998, 1999, 200, 2001; Mackey et al., 2000; Mackey et al., 2007; Rosa & Leow, 2004a, 2004b, Rosa & O’Neill, 1999, Yoshida, 2008; 2010, to name a few). Despite its increasing application, however, the use of verbal protocols in language research is not without problems (e.g., Ellis, R., 2001; Jourdenais, 2001). Two major issues with verbal protocols can have a negative impact on the internal validity of a study: reactivity and veridicality.

Reactivity refers to the impact of the data elicitation technique on the participants’ natural thought processes. For example, the cognitive demands of a verbal protocol (the need to put into words thought processes that are not normally verbalized) may alter the actual thought processes that are being observed, so that the resulting data would not reflect them accurately. Reactivity is typically a problem with concurrent verbal protocols – think-alouds – which ask the participant to speak their thoughts while performing a task. Retrospective reports, in which the original
thought processes are reconstructed rather than verbalized concurrently, are not susceptible to reactivity, since the targeted processes have already occurred at the time of the report. While these processes cannot be altered after the fact, they can still be reported inaccurately – the problem referred to as non-veridicality.

Ericsson & Simon (1993) argue that concurrent reports are likely to be a more valid instrument for capturing thought processes, because the information they contain is still fresh in the short-term memory and directly accessible for verbal reports. In contrast, retrospective protocols may not always access the information stored in short-term memory due to the time-lag between the event and the recall. Therefore, memory decay may affect the validity of the report (Egi, 2004a; Ericsson & Simon, 1993; Jourdenais, 2001).

Validity concerns, however, are not the only considerations that have to be made when selecting data elicitation instruments for classroom research. While “on-line” measures, i.e. concurrent protocols like think-alouds (see Bowles, 2010 for the most extensive overview to-date) and immediate recall (Philp, 2003) have been argued to be more effective in eliciting sound memory data (Ericsson & Simon, 1984), they are too intrusive to apply in the study of classroom communicative interaction. Since both techniques require the participants to pause and verbalize their thoughts during or immediately after the targeted event, thus halting the interaction, they would irreparably change the naturally occurring classroom communicative patterns and compromise the validity of the study.

Because the focus of the current study was on naturalistic (rather than scripted or otherwise directed by the researcher) classroom interaction, a retrospective verbal report technique known as stimulated recall interview (Gass & Mackey, 2000) was employed.
Stimulated recall protocols elicit retrospective data by presenting the participants with stimuli related to the thought processes they are trying to reconstruct. For example, in a stimulated recall interview aiming to capture decision-making processes while writing an essay, the original draft or drafts may be presented to the learner in order to “jog” their memory and help them reconstruct their thoughts at the time when they were writing the essay. Bloom (1954) argued that stimulated recall has advantage over simple post-hoc interviews because the latter rely heavily on memory without prompts, which is more susceptible to memory decay and fabrication. Stimulated recall also has the advantage of simplicity over think-aloud protocols, which require participants to be trained in verbalizing their thoughts while simultaneously carrying out a task. Ericsson and Simon (1983) and Lieberman (1979) have demonstrated that stimulated recall protocols do correspond with actual behavior, thus confirming the validity of this measure for examining and establishing features of human cognitive processes.

As discussed in the preceding paragraphs, retrospective protocols – including stimulated recall reports – are typically not susceptible to reactivity but veridicality remains a major concern in their application, necessitating special measures to ensure the collection of valid data.

Veridicality issues in retrospective reports fall under two categories: omission (largely due to memory decay) and commission (verbalizing thoughts that did not take place, typically related to fabrication, overgeneralization, or elaboration).

The problem of memory decay in the use of stimulated recall protocols can be controlled in two ways: by keeping the time lag between the event and the recall as short as possible (Bloom, 1954, suggests an optimal time-frame of 48 hours), and by selecting contextually salient cues as stimuli for the recall interviews. Stimuli that use the same channels of perception as the
original event are particularly effective. For example, in the case of a study on classroom interaction, a combination of aural and visual cues in the form of a video recording capturing the targeted events would be more likely to yield high-quality memory data than using a written report or observation notes that describe the same events, because the participants will have experienced the original interaction through hearing and seeing, rather than through reading. Presenting them with contextual cues or stimuli, employing the same perception channels (auditory-visual) is more likely to trigger accurate recall.

Fabrication is another aspect of non-veridicality in retrospective verbal protocols (Ericsson & Simon, 1984, 1993; Leow, 2002; Leow & Morgan-Short, 2004; Russo et al., 1989). It occurs for several reasons: interviewer effect, elaboration, and double exposure. Interviewer effect is argued to promote fabrication in participant’s verbal responses because the participants may be eager to “please the researchers and report behaviors that they believe would be of interest to the researcher, or behaviors they should be using as “good” learners, rather than the actual strategies and processes employed” (Jourdenais, 2001, p. 356). One way to counteract this tendency is to conduct the interviews in a non-personal, computer-assisted format: The participants are seated at individual computer stations, and the interview prompts are delivered either through the computer screen or orally, but the interviewer maintains a distance between herself and the interviewees.

Once the stimulated recall interview is underway, the researcher should refrain from reacting to participants’ comments, as well as from pushing for further comments, reasoning or elaboration (Ericsson & Simon, 1993). The provision of linguistic feedback during a stimulated recall interview should be specifically avoided at all costs, particularly in a research study focusing on awareness of corrective feedback, since the transformation of the recall interview
from a one-way information source to an interactive process can irreparably alter the participants’ recollection of their original thought processes.

Another major source of fabrication, elaboration, occurs when the participants go beyond reconstructing their original thoughts, adding to them explanations or justifications for their answers or their original behavior. The interviewee may engage in elaboration for two reasons: improperly designed prompts or questions, and inadequate probing. Questions which are too general or which ask the participants about their cognitive processes can lead to elaboration, because in such cases the participants would not base their answers on memory for specific events, but would theorize about their process (Ericsson & Simon, 1980: 245). Requests to the learners to clarify or elaborate on their responses during verbal recall sessions can also trigger fabrications because they invite “the kind of self-theorizing that can make retrospective reports nonveridical” (Russo et al., 1989).

Several measures have been suggested in the literature to help avert nonveridicality through elaboration. Firstly, prompts for recall interviews need to be carefully constructed and phrased (Gass & Mackey, 2000; Jourdenais, 2001). Prompts that are likely to elicit valid data are open-ended questions (Jourdenais, 2001), statements (rather than questions) guiding the learners’ attention to their original thoughts, e.g., “Try to recall and state what you were thinking.”, and prompts designed to elicit only attended-to information, which ask the participants simply to report their original thoughts rather than to elaborate on the reasoning behind them.

Second, the length of time allowed for verbalization also plays a role in the prevention of fabrication. Egi (2004) suggests that longer verbalizations of thoughts in recall sessions may be due to the fact that the participants “tend to explain their thoughts rather than simply verbalizing
them” (p. 255). A suitable measure to discourage fabrication, then, would be short timeframes for verbalization that would allow only for the simple statements to be produced, rather than elaborate reasoning and justifications of the learners’ thought processes. Generally, it is advisable that stimulated recall procedures not be overly structured by the researcher, to minimize the opportunities for researcher influence and interference which may compromise the veridicality of the data.

Instructions for the participants should be carefully phrased, pilot-tested, and standardized either by pre-recording them or by presenting them to all participants in written form. Standardized instructions can help orient the participants to the actual time period under recall – particularly important, since a single word or tense change in the instructions can affect the nature of the participants’ recall (Gass & Mackey, 2000). It is also important to stress in the instructions that the participants are allowed to report lack of recall, that is, it is acceptable for them to say that they do not remember or were not thinking anything during the episode under recall (Russo et al., 1989). This measure would help avoid fabrication in the case of memory decay, when participants are more prone to reporting made-up thoughts simply in order to avoid giving a “blank” report and disappointing the researcher.

Finally, double exposure can bring about fabrication because in a stimulated recall situation, the participants are exposed to the original interaction twice – the second time in the form of a contextual cue for the recall. As a result, the participants may report on their thoughts at the moment of the interview rather than recall their original thoughts (Leow, 2002; Leow & Morgan-Short, 2004; Russo et al., 1989). A related concern is the prevention of additional exposure. When additional language classes have taken place between the original exposure and the recall interview, the learners’ memories of the original event are not only susceptible to
memory decay but also likely to be influenced by information from the additional exposure to interaction and linguistic input during the consequent language classes. Therefore, it is important to schedule recall interviews not only as soon as possible after the original interaction, but in such a way as would prevent additional opportunities for linguistic input and learning to take place.

Summary

Stimulated recall protocols (Gass & Mackey, 2000) involve the provision of a prompt, cue, or stimulus, which helps the participants to recall their thought processes during a recently completed task. A visual or audio reminder of the event, such as a recording of learner interaction during the original task in the case of research on oral interaction, or a written product, in the case of research on second language writing, is presented. This reminder can jog the participant’s recall of the mental processes involved in the original event (Polio, Gass & Chapin, 2006). The accuracy of stimulated recall is significantly enhanced when the interviews are conducted close to the event being recalled.

A major issue with the potential to undermine the validity of the data obtained by means of a stimulated recall interview, is nonveridicality. Nonveridicality presents itself in two major ways – memory decay, and memory fabrication. A number of steps can be taken to prevent the data from being contaminated and its validity compromised, including the timing of the recall interview, the selection of prompts, the formulation of instructions, and the interview procedure. Table 1 summarizes these measures.
<table>
<thead>
<tr>
<th>Type of non-veridicality</th>
<th>Source of non-veridicality</th>
<th>Preventative measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory decay</td>
<td>Time-lag between original event and recall</td>
<td>Recall should be as immediate as possible (recommended time-lag is less than 48 hours).&lt;br&gt;Contextually and perceptually salient cues should be used as stimuli.</td>
</tr>
<tr>
<td>Fabrication</td>
<td>Interviewer effect</td>
<td>Interview should be one-way information channel, rather than interactive.&lt;br&gt;“I don’t remember” reports are allowed.</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Carefully selected prompts:&lt;br&gt;Open-ended;&lt;br&gt;In the past tense;&lt;br&gt;Asking about event, not process;&lt;br&gt;No requests to clarify answers.</td>
<td>Time allowed for response should be short.</td>
</tr>
<tr>
<td>Double exposure</td>
<td>No additional exposure between event and recall allowed.</td>
<td>Prompts should be unambiguous and worded in the past tense.</td>
</tr>
</tbody>
</table>
Effectiveness of corrective feedback in language learning

The effectiveness of corrective feedback in second language learning has been examined in both laboratory and classroom settings (see Mackey, 2006, for review). Findings from individual studies, synthesized by meta-analyses, so far have demonstrated a facilitative effect for corrective feedback on L2 development (Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007; Russel & Spada, 2006). As a result, corrective feedback is now considered to be beneficial for language learning by most researchers in the field of SLA. At present, the focus of the feedback research community has shifted from studying the role of feedback in general, to examining the benefits of individual types of feedback, as well as different configurations of context, learner characteristics, and interaction, in order to gain insight into the specific processes that enable or constrain the impact of feedback on language learning.

Attention and awareness are two processes that have been hypothesized to influence the potential effect of corrective feedback. Theoretical and empirical research has highlighted the importance of conscious “selective attention” (Long, 1996) for perceiving the corrective intent of the teacher’s feedback, and for pinpointing the locus of the error. These two conditions – that feedback is understood as such and that the error is correctly identified, have been argued to determine the usability of feedback (Carroll, 2001; Gass, 1997; Gass & Varonis, 1994; Schmidt, 1990).

One way to reveal how corrective feedback facilitates L2 development is through examining how L2 learners perceive corrective feedback supplied to them during L2 interaction. Learners in L2 classrooms each have their own unique interactions with the linguistic input
provided to them by their teachers and classmates. Research into learner perceptions about corrective feedback (Egi, 2008; Gass & Lewis, 2007; Kim & Han, 2007; Mackey, Gass, & McDonough, 2000; Nabei & Swain, 2002; Roberts, 1995) is increasing.

Initial studies on learners’ perceptions in L2 classroom contexts (Allwright, 1984; Slimani, 1989), found that learners’ reports about classroom routines were idiosyncratic and that learners’ perceptions about the same classroom events differed considerably. The first study to actually examine learners’ perceptions about corrective feedback in an SLA context was Roberts (1995). Roberts investigated the extent to which college learners noticed teacher-provided error correction in a beginning-level Japanese language class. Results showed that of 92 total instances of error correction, the students were able to identify 32 on average (35%) and understood about 19 (21%). Roberts hypothesized that the efficacy of error correction is not only related to students’ perceptions about corrections, but also to the understanding of the nature of those corrections, including target of the feedback and the type of feedback. Subsequent studies utilized more refined definitions in order to explore further what impacts learners’ perceptions about corrective feedback. Several factors have been hypothesized to play a role in learner awareness of feedback, including the type of the feedback, the linguistic feature that triggered the correction, and the direction of the feedback.

Linguistic target

Roberts (1995) found that the learners in his study recognized corrective feedback more accurately when it targeted phonological errors, and least accurately if it was triggered by syntactic errors. The finding that the noticeability of the feedback is influenced by the linguistic
target of the feedback was further supported by Mackey, Gass, and McDonough (2000). In their laboratory study, en learners of English as a second language (ESL) and seven learners of Italian as a foreign language (IFL) participated in dyadic interactional tasks with native speakers and were provided with corrective feedback in the form of negotiations and recasts. Immediately after the completion of the tasks, each learner viewed the recorded interaction and participated in a stimulated recall interview (see Gass & Mackey, 2000, for further details about stimulated recall). This study found that learners were most accurate in their perceptions about lexical and phonological feedback, and much less accurate in their recognition of morphosyntactic feedback. Morphosyntactic feedback was often perceived as pertaining to semantics for the ESL learners and lexis for the IFL learners. Mackey et al. (2000) proposed that inaccurate perceptions about morphosyntactic feedback stemmed from the differing degrees of salience of morphosyntax and phonology – morphology often does not interfere with overall understanding in the same manner as incorrect pronunciation or inaccurate lexical usage.

In a conceptual replication of Mackey et al. (2000), Gass and Lewis (2007) examined Italian heritage and non-heritage learners’ perceptions about corrective feedback. Their results confirmed Mackey et al.’s (2000) findings that both non-heritage language learners and heritage language learners perceived phonological and lexical feedback much more accurately than morphosyntactic feedback. However, the two groups differed in regard to their recognition of semantic feedback, with heritage learners much more likely to interpret this feedback target correctly.

Kim and Han (2007) also found a significant relationship between students’ perceptions about corrective feedback and the type of linguistic target. Four intermediate English as a foreign language (EFL) classes at a private institute in Seoul Korea were videotaped and participated in a
stimulated recall protocol was conducted. The results showed that phonological feedback in the observed classes was perceived more accurately than lexical feedback, which in turn was perceived more accurately than morphological corrections. Kim & Han’s (2007) suggested that the lower frequency of noticing of morphosyntactic feedback in complex recasts may have been driven by learners’ natural tendency to process the meaning of the input first (VanPatten, 1996, 2004).

The results from an Arabic language classroom study (Mackey et al, 2007), however, are not consistent with the findings described above. In a study of two intact beginning-level Arabic language classes, Mackey et al found that learners of Arabic were more likely to perceive accurately corrective feedback that targeted morphology or lexis, than to recognize phonology. This finding may have been the result of the specific coding scheme that Mackey et al used, in which the categories of lexis and morphology were collapsed into one category – morpholexis, which was felt to describe better the linguistic reality of the Arabic language.

These inconsistent findings highlight the complex relationship between learner recognition of corrective feedback, and its linguistic target. It is important to note that researchers have argued that not all linguistic forms may be processed and acquired in the same manner (DeKeyser, 2005; Long, 2006; VanPatten, 1996). However, we do not yet know enough about the processes behind these differences. More research is needed into the relationship between the linguistic focus of the corrective feedback, and the way learners engage with it.
Type of feedback

The type of the corrective feedback was also found to be related to the learners’ recognition of feedback in Roberts’ (1995) study: Learners were more likely to report awareness of feedback that been given in the form of recasts, than if it had been delivered as negotiation or elicitation.

Kim and Han (2007) also investigated the relationship between students’ perceptions about the linguistic target of corrective feedback in relation to different types of recasts. They found that learners perceived teachers’ corrective intentions when corrective feedback was provided through declarative recasts more often than when interrogative recasts were used. They argued that “an interrogative recast is more implicit as corrective feedback since the learner may interpret it as either corrective or as a request to confirm the intended meaning.”

Mackey et al. (2007) found that beginning learners of Arabic were more likely to recall accurately the linguistic target of a correction, if it had been delivered in explicit form, rather than in the form of a declarative recast or negotiation. However, they also noted that interrogative recasts – an implicit type of feedback, were perceived correctly almost as frequently as explicit feedback.

In addition to the target of feedback and the type of feedback, research has also examined the effects of the direction of the feedback on learners’ awareness of corrective feedback.
Direction of the feedback

The nature of a learner’s engagement with a feedback episode, i.e., whether or not the correction was directed at them or at someone else in the classroom, can also influence the extent to which the learner becomes aware of the correction. In a case study, Nabei and Swain (2002) discovered that their participant as a rule did not attend to feedback if it was targeted towards other learners. However, Ohta (2000) investigated learner reactions to classroom feedback, including silent speech, and reported that, in fact, beginning and intermediate learners of Japanese were equally likely to react to recasts that were not addressed to them. In Kim and Han (2007)’s study, no difference was found in the accuracy of perceptions between direct or indirect addressees of the feedback. Mackey et al. (2007) reported that their beginning Arabic learners were significantly more likely to perceive the linguistic target of the corrective feedback, if it was directed at them. Due to these contradictory findings, the relationship between the nature of learner participation and learners’ perceptions about the linguistic corrective feedback is still unclear.

Learner proficiency

Developmental studies have shown that learners at different proficiency levels benefit differently from corrective feedback (Ammar & Spada, 2006; DeKeyser, 1993; Havranek, 2001; Havranek & Cesnik, 2002; Mackey & Philp, 1998, Trofimovich, Ammar & Gatbonton, 2007). The underlying reason for the differing developmental results has often been assumed to be the learners’ differing abilities to notice and process input, particularly feedback, due to their differing exposure, familiarity and increased automaticity as a result of repeated practice, which
allow more attentional resources to be allocated to the other cognitive stimuli, in addition to the ones immediately at hand (Philp, 2003).

However, few studies have addressed empirically the question of whether learners at different developmental or proficiency levels indeed show different levels of feedback recognition.

In a small-scale classroom-based study, Roberts (1995) determined proficiency on the basis of the instructor’s judgment of the relative linguistic mastery of the three participants in his study. Findings from his small-scale study show that the highest proficiency learner recognized corrective feedback less frequently than both the medium and the low proficiency learners. Roberts hypothesized that the inverse relationship between proficiency and noticing in his study may be due to the differing amounts of exposure to error-correction. Lower-proficiency learners are more likely to make errors and thus more likely to be the recipients of larger quantities of corrective feedback, hence they may be “more finely tuned to the teacher’s error correction signals” (Roberts, 1995, p. 179).

In a larger, laboratory study, Philp (2003) investigated the frequency and accuracy of noticing among 33 ESL learners at three different developmental levels: low, intermediate, and high. The developmental levels of the learners were determined on the basis of their developmental stages in relation to question formation (Pienemann, 1995). Results from her study indicate that higher-level learners recalled recasts more accurately than lower level learners, a finding contrary to Roberts’ result.

Lin & Hedgcock (1996) also investigated the influence of proficiency on learner noticing of correction. In a laboratory study of 8 Chinese L1 learners of Spanish as a second language,
they found that the four higher proficiency learners, who were at a very advanced level of language mastery, received much fewer corrections than the lower-proficiency learners, and incorporated the corrections 9 times more often than lower-proficiency learners.

Finally, Trofimovich et al. (2007) examined the role of proficiency in learner noticing of recasts in a computer-based interaction task. They operationalized proficiency in terms of demonstrated quantity and rate of speech and grammatical accuracy and found that proficiency was not a factor associated with noticing of recasts but was a factor in language development with higher-proficiency learners benefitting more from recasts than lower-proficiency learners.

To summarize, findings about the role of proficiency in the noticing of corrective feedback are few and conflicting. Only one very small scale study has been conducted so far in a classroom, and no studies have explored proficiency in the context of classrooms at different points on the proficiency continuum, rather than learner relative proficiency within a homogenous group. To address this gap, the current study will look at beginning and advanced classrooms defined in terms of instructional or institutional levels.
CHAPTER 3: DESIGN AND METHODS

Introduction

The current study draws from research on interactional feedback and awareness in order to explore the ways learners of Arabic perceive classroom feedback, and the factors related to their perceptions. Since research on the topic of Arabic classroom instruction in general, and on the role of classroom interaction and feedback in the formal instruction of Arabic in particular, is currently scarce, a broader exploratory approach was deemed necessary to allow a fuller picture of the phenomenon at hand to emerge. Hence, a mixed-analysis descriptive design was selected for the study, rather than a statistically more robust but also more restrictive experimental design.

This chapter starts with a summary of the relevant issues and questions that emerged in the literature review. On the basis of this summary, I then formulate the research questions and hypotheses that guided the study, and describe the methodology employed to address them.

Summary of research

In the previous chapter, constructs relevant to the current study were introduced and discussed. To summarize, interactional feedback has been shown to play a facilitative role in language development (see reviews in Li, 2010; Mackey & Goo, 2007; Russell & Spada, 2006). While the effect of feedback on language learning has been extensively documented, the mental processes underlying this effect remain unclear. The preponderance of studies addressing this gap has focused on the cognitive aspects of feedback processing, particularly on attention,
perception, and noticing. Several factors that may impact the extent to which feedback is noticed in the input have been identified, including characteristics of the learning situation (grammar-based vs meaning-based classroom), characteristics of the data-collection situation (lab vs classroom), characteristics of the feedback (type, target, and direction), and characteristics of the learner (e.g., proficiency).

While these factors have been shown to play a role in language development in response to corrective feedback, the relationship between these factors and the learners’ awareness of the corrective feedback they receive remains understudied. The current study was designed to address the described gap, by exploring the following five research questions.

Research Questions and Hypotheses

RQ 1: Are learners aware of a) the corrective feedback, occurring in their classroom during oral interactive activities, and, b) the linguistic targets of this corrective feedback?

Hypothesis: Previous research has yielded conflicting results regarding learner awareness of corrective feedback, ranging from very low (Roberts, 1995) to very high (Mackey, 2002; Philp, 2003). Research suggests that learners may identify corrective feedback more easily in form-focused environments, compared to meaning-focused environments (Lyster & Mori, 2006). Because the current study was conducted in language classrooms with distinct orientation to form, it was hypothesized that the learners would show heightened awareness of both correction and linguistic targets of feedback.

RQ2: Is there a relationship between the learners’ proficiency level and their awareness
of a) the corrective intent of the feedback, and b) the linguistic target of the feedback?

*Hypothesis:* Laboratory research on uptake and language development suggests a positive relationship between the learner’s developmental or proficiency level and their ability to benefit from feedback. Researchers have hypothesized that the cognitive mechanism underpinning this relationship may be one of noticing and awareness increasing with proficiency (Ammar & Spada, 2006; Lin & Hedgcock, 1996; Philp, 2003). On this basis, advanced learners in this study were expected to report awareness of both correction and linguistic target more frequently than beginning learners.

**RQ3:** Is there a relationship between the characteristics of the feedback (type, linguistic target, direction) and the learners’ awareness of correction?

*Hypotheses:* On the basis of existing empirical research, it was predicted that certain characteristics of the feedback – explicitness of the feedback move and salience of the linguistic target - will be associated with higher rates of reported awareness. More explicit feedback was expected to be perceived accurately more often than implicit feedback. An association between the linguistic target and the rate of awareness was expected as well, with the corrective intent of lexical episodes more accurately perceived than that of morphological, phonological and syntactic episodes (Kim & Han, 2007; Mackey et al., 2000; Roberts, 1995). Since findings on the relationship between the direction of the feedback and learner awareness are inconclusive (Kim & Han, 2007; Mackey et al., 2007; Nabei & Swain, 2004; Ohta, 2000)(Gass & Lewis, 2007), no hypothesis could be formulated regarding this variable.

**RQ4:** Is there a relationship between the characteristics of the feedback (type, linguistic target, direction) and the learners’ awareness of the linguistic target of the
correction?

_Hypotheses:_ Similarly to the previous research question, a positive association was expected between the explicitness of feedback and learner awareness of its linguistic focus (Roberts, 1995; Mackey et al., 2007). Certain types of linguistic targets were also expected to be more readily identifiable than others – e.g., lexical feedback was expected to be more accurately perceived than morphological or syntactic (Gass & Lewis, 2007; Mackey et al., 2000). The direction of the feedback was also hypothesized to influence the rate of reported awareness (Mackey et al., 2007) but a hypothesis about the nature of this influence could not be formulated based on the existing empirical research.

In order to explore further the factors that influence learner awareness of corrective feedback, a qualitative analysis was carried out in addition to the quantitative one. The qualitative analysis was guided by the following question:

RQ5. What factors do learners report as influencing their awareness of corrective feedback?

Because this was an exploratory qualitative question, an inductive approach (thematic analysis) was selected as a framework for the data analysis. Inductive analysis yields best results when it is applied free of restraints such as pre-determined coding categories (Mackey & Gass, 2005). Hence, to preserve the integrity of the data analysis method, no hypothesis was formulated in advance.

The variables posited in these research questions will now be described in more detail. This study investigated the relationships between learners’ awareness of the corrective intent and linguistic target of a corrective feedback episode on the one hand, and characteristics of the feedback episode (type, linguistic target, direction) and of the learners (proficiency), on the other.
Figure 1 presents the independent and dependent variables under study, and the hypothesized relationships among them.

![Diagram showing independent variables, mediating variable, and dependent variables]

**Figure 1. Study variables.**

**Operationalizations**

The following variables were investigated in this study: characteristics of the corrective feedback, which included type, target, and direction of the feedback, learner proficiency, and learner awareness of the corrective intent or the linguistic target of the corrective feedback. Details about each category, together with illustrative examples, will be provided in the coding section of this chapter.
Corrective feedback

A corrective feedback episode was defined as a communicative exchange between a learner and the instructor, in which the learner produced a non-targetlike utterance, the teacher provided corrective feedback that addressed one or more linguistic problems in the learner’s utterance, and the learner either responded to the teacher’s feedback in some way, or continued without response\(^1\). Example 5 shows a typical corrective feedback episode from the data\(^2\):

Example 5. Feedback episode.

L: *mundhu waaHida sana* (Syntactic error – word order)
   (since *one year)
   A year ago

T: *sana waaHida* (Declarative recast)
   (year one)
   One year

L: *sana waaHida* (Response – Repair)
   (year one)
   One year

Proficiency

Proficiency in this study was operationalized in terms of institutional or instructional levels (Thomas, 1994), i.e., the proficiency of the participants was determined on the basis of the level of the class that they were taking – beginning or advanced. Learners were defined as having

\(^1\) The types of learner reactions to corrective feedback – referred to as uptake - and their relationship to the learner’s motivation, thought processes, and diverse learning outcomes, have been examined extensively within the interactionist literature (e.g., Ellis, Basturkmen & Loewen, 2001; Loewen, 2005; Loewen & Philp, 2006; Lyster, 1998; Lyster & Ranta, 1997; Mackey & Philp, 1998; Panova & Lyster, 2002; Sheen, 2004, etc.). While this topic is an important piece of the puzzle in the exploration of the learners’ cognitive engagement with corrective feedback, it falls beyond the scope of the current study and will not be addressed in this dissertation.

\(^2\) To preserve the emic perspective of the analysis, in all examples throughout this text learner errors are defined on the basis of the instructor’s corrective intentions, explicitly stated in the instructor’s stimulated recall comments. In cases when the learner’s utterance contained non-targetlike forms that were not addressed by the instructor’s feedback, nor recalled in the instructor’s comments, such forms were not marked as errors (with an asterisk).
beginning level proficiency if they were enrolled in a first-year Arabic language course, and
advanced level proficiency if enrolled in a third year Arabic course. This section presents a
rationale for the selection of this approach to operationalizing proficiency within the current
contextual constraints of the Arabic language teaching field.

Traditionally, in the empirical research on corrective feedback and language
development, participant proficiency has been measured through the use of standardized tests of
general proficiency (e.g., TOEFL), developmental levels (e.g., Pienemann, 1988; 1998; 2003;
2005; Pienemann & Johnston, 1987; Pienemann, Johnston & Brindley, 1988, etc.), or tailor-made
tests (e.g., Loewen, 2005). At present, these approaches have somewhat limited applicability to
the study of Arabic language learning for a number of reasons, which will be explained below.

Standardized tests, a proficiency measure frequently used in SLA research, are difficult to
come by in the field of Arabic learning. Winke & Aquil (2006) list three main reasons for the
scarcity of such tests: The small numbers of learners enrolled in Arabic language courses do not
allow for necessary test item callibration and norming; the small sizes and the predominantly
teaching focus of the current Arabic language programs make the development of in-house
standardized tests difficult; finally, the diglossic nature of Arabic complicates the specifications
for such tests, because it is difficult to determine in which variety language functions should be
carried out. The authors contend that as a result of these issues the existing standardized tests are
not only few and far between, but often limited in scope – they are used to test a single skill or a
small number of skills, for example, reading or speaking and listening, or reading and writing,
rather than to determine overall proficiency across all language skills (Winke & Aquil, 2006).

Developmental levels, or systematized descriptors of the state of the learner’s
interlanguage, based on Pienemann’s Processability Theory (1998, 2003) have been widely used
in developmental studies of corrective feedback (e.g., Mackey 1999; Mackey & Philp, 1998; Philp, 2003, etc.) and have been accepted as an efficient and reliable measurement of proficiency. However, research on the stages of interlanguage development in the study of Arabic as a foreign language has been very limited in both numbers and scope, and the results from this research have been inconclusive (AlHawary, 1999; 2009; Mansouri, 2005; Nielsen, 1997). Until a hierarchy of developmental stages has been established for Arabic, the reliability of this approach to determining the proficiency of learners of Arabic remains questionable.

Tailored tests allow proficiency to be measured in relation to a specific linguistic structure or number of structures. This is a commonly applied measurement in experimental studies (Mackey & Gass, 2005). The use of such tailored tests, however, is only possible when the structures that will be the focus of the study are known in advance or their occurrence and frequency can be manipulated as part of the research design. The nature of the current study precluded the use of this measure. The object of this study was naturally occurring classroom feedback, rather than scripted feedback. This type of feedback arises – by definition – as a reaction to the learners’ actual non-targetlike utterances during the observation period, and is not limited to prescribed feedback types or linguistic targets. Since it is impossible to predict which linguistic structures may become the focus of the feedback during an authentic classroom exchange, establishing proficiency in relation to pre-determined linguistic structures was not a feasible solution in this study (see Loewen, 2005, for a similar argument). Using naturally emerging, non-scripted data yields results that are statistically less robust, and this limitation of the current study should be considered when interpreting the results. However, focusing on this type of data also increases the ecological and pedagogical validity of the results which was an important consideration in this exploratory classroom study.
The remaining alternative for operationalizing the proficiency of the participating learners, therefore, was in terms of institutional status (Thomas, 1994; 2006) - that is, according to the proficiency levels assigned to them by the university. This approach to defining proficiency is currently the most commonly used one in SLA studies due to its convenience and efficiency (Ortega & Byrnes, 2008; Thomas 2006). In this study, institutional status was deemed a suitable and sufficiently sensitive measure for two reasons. First, the two groups selected for data collection were at opposite points on the proficiency continuum, namely at the very beginning and the very end of the language program where the data collection took place; the juxtaposition of these distinct positions highlighted maximally the existing differences between the two groups. Second, the study was descriptive rather than experimental in nature and language development in relation to particular linguistic structure - which would have necessitated establishing proficiency relative to those structures - was not the focus of inquiry (Thomas, 1994). On this basis, the participants' institutional level was taken to be a sufficient indicator of their proficiency, and a suitable way to operationalize this variable. Thus, beginning and advanced learners in this study were learners enrolled in the first and fifth semester of Arabic, respectively. Throughout the study, learners at the beginning level of proficiency are referred to as “beginners”, “beginning learners” or “lower proficiency learners/participants”. Similarly, learners at the advanced level of proficiency are referred to as “advanced”, “advanced learners”, and “higher proficiency learners/participants”.

Assessment

Following Allport (1988), awareness was defined in terms of the participant’s ability to recall the events that were the focus of the study. The assumption made in selecting this
definition is that such recall constitutes a reliable representation of the participant’s actual thought processes at the time of the original event. Awareness of correction was operationalized as verbal comments in the participants’ recall protocols indicating that the learners had noticed either the corrective intent of the teacher’s intervention, or the presence of an error, or both. Awareness of linguistic target was similarly operationalized as the explicit mention of the linguistic structure which had been the focus of the teacher’s intervention. Throughout the study, the following terms are used synonymously with awareness: “to be aware of X”, “to notice X”, “to attend to X”, “to direct attention to X”, and “to perceive (accurately) X”.

The data collection procedure will now be described.

Design and Procedure

An ex post facto mixed-method (quantitative-dominant) design was employed to address the research questions. This is a descriptive type of design, characterized by the integration of quantitative and qualitative analyses of the data in which the qualitative analysis plays a supportive role by providing additional context for interpreting the quantitative findings (Teddie & Tashakkori, 2009). Data collection took place over a period of four semesters in a large undergraduate-level Arabic FL program at a private US university. Learners and teachers in four beginning and one advanced language classrooms were observed and videotaped during authentic classroom interaction, and subsequently interviewed using a stimulated recall interview procedure. Following the interviews, the learners also filled out a questionnaire about their

---

3 The number of participating classrooms at each level was limited by logistical and practical constraints (see a detailed list on page 67): While the low number of participants, particularly at the advanced level, is an obvious limitation to this study, it still allowed for statistical tests of the data to be carried out and conclusions to be drawn. Subsequent studies with larger participant pools are needed for a clearer description and a better understanding of classroom interaction processes.
learning practices and attitudes to corrective feedback. The following sections will provide a more detailed description of the participants, context, materials, and data collection procedures.

Figure 2 gives an overview of the data collection procedure.

*Figure 2. Procedure.*
Context

The data collection took place in a large undergraduate-level Arabic language program at a private US university. At the time of data collection, the program comprised three levels of Arabic language instruction: beginning (semesters I and II), intermediate (semesters III and IV), and advanced (semesters V, VI, and beyond). Instruction at the beginning and intermediate levels focused on the development of general language capabilities in speaking, listening, reading and writing MSA, as well as a general introduction to the culture of the contemporary Arab world. Teaching at the advanced level was differentiated into several instructional foci: continued development of general language capabilities, specialized instruction in FSA, and content courses taught in MSA (e.g. courses on contemporary Arabic media, geographical and sociopolitical survey of the contemporary Arab world, survey of classical and contemporary Arabic literary genres and texts).

Since instructional foci and classroom content differed widely at the advanced level, only general language classrooms were selected for participation in the study, in order to establish a basis for comparison between interactional patterns and behaviors. While each of the four participating instructors had a unique teaching style and interactional preferences, they all based their instruction on identical curricular guidelines and achievement expectations, and applied the same proficiency-based, structurally oriented approach, outlined in the textbook series used in the general language classrooms – *Al-kitaab fii ta'allum al-'arabiyya* (Brustad, Al-Batal & Al-Tonsi, 2004; 2006).

Typical interactional patterns in the participating classrooms included teacher-led whole class discussions, small-group work, teacher- and learner- initiated whole class question- and-answer activities, teacher-elicited individual learner participation, and spontaneous self-initiated
learner participation. The interactive classroom activities at the beginning level were almost exclusively textbook-based, and revolved around discussions of the instructional texts and audio-video materials or focus-on-forms teacher- or learner-initiated discussions. At the advanced level, the classroom activities deviated from the textbook more frequently, with the teacher relying more heavily on supplementary reading and audio-video materials, and spontaneously occurring or learner-initiated topics for discussion. Still, at this level, too, the linguistic content was kept as close as possible to the one provided in the textbook and whole class or individual drilling, testing or practice of structures and forms were a typical occurrence during the observational period for this study.

The context for the present study, therefore, can be characterized as largely forms-focused, rather than meaning-focused. In this respect, it differs in significant ways from the typically communicative, conversation-based, fluency-oriented second language classrooms in which much of the research on corrective feedback perceptions and efficacy has been conducted. Such differences may have implications for the ways learners perceive, interpret, and respond to corrective feedback, as suggested by previous research (Kim & Han, 2007; Lyster & Mori, 2006).

Specifically, learners in predominantly forms-focused instructional environments may be more attuned to corrective feedback that targets overtly grammatical forms, especially morphology and syntax, because this is the type of feedback that they would most often receive. Such learners would be more likely to focus their attention actively on the structural corrections in the instructor's feedback, rather than on its overall meaning or message. As a result, learners in a forms-focused environment would be likely to notice the instructor's recast of a non-targetlike morphological form, for instance, partially because they have the expectation that the instructor's
role in the conversation is strictly to validate or sanction their language production. Learners in more meaning-focused classrooms, on the other hand, may be more accustomed to receiving feedback that targets the general meaning or comprehensibility of their utterances, rather than specific structural features. They would also be more likely to view the instructor as an equal participant in the conversation, and his or her utterances – as genuine response to the learner's contributions. Thus, they may be more prone to overlook the corrective intent of the feedback they receive and less likely to notice the correction, realize the gap between their own utterance and the correct form, and incorporate the new knowledge into their interlanguage.

Consequently, findings regarding the ways learners perceive corrective feedback in forms-focused environments should not be assumed to extend to learners in meaning-based classrooms as well. The results from the current study should be interpreted with this caveat in mind.

Observations and participant selection

As a preliminary stage to the data-collection process, the researcher observed 10 language classrooms – 7 beginning and 3 advanced⁴, in order to identify appropriate contexts for collecting classroom data and recruiting participants. From the initial 10 classrooms, 4 beginning and 1 advanced were eventually selected on the basis of the following criteria: a) the instructional approach included frequent teacher-led whole class discussions with ample learner participation; b) the instructor frequently provided corrective feedback in response to learners’ non-target like utterances; c) the instructors and learners agreed to participate in the study and be

---

⁴ The data were collected over four semesters. The number of initially observed classrooms reflects the total number of available Arabic language sections during those four semesters whose instructors agreed to be observed, within certain scheduling constraints – i.e. in cases when multiple sections took place at the same time, only one of them could be observed.
videotaped during classroom interaction, and d) the class was scheduled early enough in the day to allow for same-day stimulated recall interviews to take place.

Each of the participating language classrooms was observed for a period of time that allowed for ten class meetings to take place. Because of differing course schedules and syllabi, the length of this observational stage was different for each class and proficiency level, ranging from two-three weeks for the beginning classes, to seven weeks for the advanced class. During the observational stage, the researcher collected data about the classroom interaction patterns and occurrences of corrective feedback through extensive field notes, which were occasionally supplemented with audio- and video-recordings. The data from the observational stage served to create the corrective feedback profile of the participating classes.

**Participants**

Data were collected from five intact Arabic classes - four beginning and one advanced - with a total of 73 enrolled learners. The beginning classes met for 65 minutes 5 times a week and the advanced class took place twice a week for 75 minutes. During the observational stage of the study, the number of learners present in the observed classes ranged between 7 and 17 (M=12). At the time of classroom video-recording for interview purposes, class sizes ranged between 11 and 15 learners (M=13). Participants at the stimulated recall interview stage of the study were 31 learners from these classes, 26 beginning and 5 advanced, and their four instructors (one instructor taught two beginning classes). A description of each group is given below. All interviewed participants received a $10 gift certificate as compensation for their time.
**Beginning learners**

Twenty-six learners from the four observed beginning-level classrooms volunteered to participate in the stimulated recall interview. All were native speakers of English. Two learners indicated in the background questionnaire that they were bilingual or multilingual – in addition to English, they listed Chinese, Uigur, Russian and Kirgiz as their first or domestic languages. Learners who had indicated that they were heritage learners of Arabic were excluded from this stage of the study, since existing research suggests perceptual and processing differences between heritage and non-heritage beginning learners in relation to corrective feedback (Gass & Lewis, 2007). Arabic was a second or third foreign language to the majority of the learners in this group (n=24). Participants indicated most often that they had previously had formal instruction in French (n=11), Spanish (n=11), German (n=5), or Hebrew (n=3), and less frequently in Italian, Polish, Romanian, Russian, Thai, and Uzbek (n=1 each).

Participants’ ages ranged between 18 and 22, with a mean of 19. There were 15 female and 11 male learners in this group. At the time of data collection, the learners were either in the final week of their first semester (Classes 1 and 3), or in the third week of their second semester of study (Classes 2 and 4), and had received either 3 or 4 months of Arabic language instruction, respectively (mean 3.4). On the basis of the learners’ classroom behavior and capabilities, estimated during the observational stage of the study, this amount of language instruction was deemed sufficient to equip the learners for participation in a meaning-based conversational activity.

**Advanced learners**

Five advanced learners from the same intact class participated in the study – three female
and two male. All were of English language L1 background, with ages ranging between 20 and 21, mean 20.4. At the time of data collection, all learners were enrolled in their fifth semester of Arabic. Four of the learners had taken Arabic courses only at their current Arabic language program, and one had completed two semesters in a study-abroad program in Yemen. For all participants Arabic was a second or third foreign language; other languages in the participants’ backgrounds included Spanish (n=3), German (n=2), and French (n=1).

Instructors

Three instructors volunteered to participate in the study together with the four classes they were teaching at the time of data collection (one instructor was teaching two beginning-level classes). The fourth instructor was the researcher. All beginning-level instructors were female. One of them was a native speaker of Arabic, and the other two – nonnative speakers (L1 English and Bulgarian). The advanced-level instructor was a male native speaker of Arabic. All instructors had eight or more years of experience in teaching Arabic as a foreign language at the college level, and three or more years of teaching in the language program where the data collection took place.

Informed consent procedure

Before video- or audio-recording took place, all 73 learners from the observed classrooms, as well as the four participating teachers, signed informed consent forms which contained information about the study goals, procedures, participants’ rights, and the measures taken to ensure participant anonymity. A copy of the consent form is provided in Appendix B.
Background questionnaire

After the stimulated recall interview, all participants completed a 17-item questionnaire. The questionnaire was intended to gather data about the participants’ age, gender, first language(s), additional language studies, Arabic language learning experience, dialect exposure, motivation for studying Arabic, attitudes to interacting in Arabic, and attitudes to teacher-provided classroom feedback. A list of the questions is provided in Appendix C. The collected information was used to create profile descriptions of the two groups of participants (see Participants).

Classroom interaction activities

Findings from previous research suggest that the frequency, characteristics, and noticeability of corrective feedback may be influenced by the type of classroom activity in which the learners engage (Oliver & Mackey, 2003). For instance, learners are most likely to receive corrective feedback and attend to it during instructional, particularly forms-focused activities (e.g., a grammar review, a vocabulary game, etc.) and least likely to be given feedback and to pay attention to it during administrative tasks or meaning-based activities (e.g., warm-up conversation and “housekeeping” activities like checking attendance, making announcements or assigning homework). To provide basis for comparison between the five participating classes in this study, care was taken to select for videotaping only classroom activities which elicited similar amounts and patterns of interaction, and provided similar opportunities for the provision of corrective feedback.

The selected activities ranged from a semi-authentic communicative task to more artificial textbook-based exercises, but had similar instructional goals and resulted in similar
amounts of linguistic output and teacher-provided corrective feedback. Table 2. provides information about the participating classes and interactional activities.

Table 2. *Class information*

<table>
<thead>
<tr>
<th>Class</th>
<th>Teacher</th>
<th>Level</th>
<th>N learners</th>
<th>Activity and format of the interaction</th>
<th>N episodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>T1</td>
<td>Beginning</td>
<td>11 (6)</td>
<td>Comprehension questions on the textbook video: Daily routines</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Format: Teacher-learner</td>
<td></td>
</tr>
<tr>
<td>Class 2</td>
<td>T2</td>
<td>Beginning</td>
<td>12 (6)</td>
<td>Information-gap survey: choosing a roommate</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Format: Learner-Learner</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Teacher-Learner</td>
<td></td>
</tr>
<tr>
<td>Class 3</td>
<td>T3</td>
<td>Beginning</td>
<td>14 (7)</td>
<td>Asking and answering questions: Practicing new vocabulary</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Format: Teacher-Learner</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Learner-Learner</td>
<td></td>
</tr>
<tr>
<td>Class 4</td>
<td>T3</td>
<td>Beginning</td>
<td>15 (7)</td>
<td>Asking and answering questions: Practicing new vocabulary</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Format: Teacher-Learner</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Learner-Learner</td>
<td></td>
</tr>
<tr>
<td>Class 5</td>
<td>T4</td>
<td>Advanced</td>
<td>13 (5)</td>
<td>Discussion of textbook reading: Generation gap</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Format: Teacher-Learner</td>
<td></td>
</tr>
</tbody>
</table>

Note: Classes 3 and 4 were taught by the same instructor, Teacher C. N learners = the number of learners present in class at the time of video-recording, followed by the number of learners who volunteered for stimulated recall interviews (in brackets). N episodes = total number of episodes from each class (both corrective feedback and distractors), used for stimulated recall interviews.
Class 1 (beginning): The learners and their teacher participated in a survey communication task – all learners were given identical sets of questions, to which they had to find answers by interviewing several of their classmates as well as the teacher. The topic of the survey was finding a new roommate: the learners were asking and answering questions about individual habits, routines and preferences, in order to find a suitable person to share their apartment. The teacher interacted with the learners as a regular participant in the task, while also monitoring the work of individual pairs and providing corrective feedback whenever needed. The activity ended with a teacher-fronted whole-class recapitulation session during which individual learners reported on the information they had collected. The teacher continued to provide corrective feedback for learner non-targetlike utterances at this stage of the activity.

Class 2 (beginning): The whole class watched a videorecording from the course instructional packet. In the video, a familiar character described his daily routine. After the viewing, the teacher asked individual students comprehension questions on the content of the videorecording, as well as individualized questions about the learners' own daily routines, and provided corrective feedback for learner non-targetlike utterances whenever appropriate.

Classes 3 & 4 (beginning): Both classes were observed during the same type of activity, which was based on a textbook exercise aimed at vocabulary and grammar review. The learners were given identical lists of questions in English, which they had to translate into Arabic before addressing them to their classmates. They were allowed first to work in groups and help each other in formulating the questions and answers. At this stage the teacher monitored the group work and provided information and corrective feedback as needed. The learners were then brought together into a whole-class teacher-fronted recapitulation session in which the learners took turns asking and answering questions in pairs. The teacher provided corrective feedback to
most learner non-targetlike utterances at this stage.

Class 5 (advanced): The learners had come to class prepared to discuss a short story from the course textbook. Together with the teacher, the class discussed both the contents of the story, and their own reaction to it. Learners answered comprehension questions, asked by the teacher, as well as commented on their classmates’ contributions to the discussion. The teacher provided corrective feedback to learner non-targetlike production whenever needed and/or appropriate.

Data collection procedure

The data collection proceeded as follows: One class meeting for each participating class was conducted in a specially equipped observational classroom. Visual and audio-signal from the classroom was simultaneously transmitted via two remotely controlled videocameras positioned to capture the activity in the whole classroom to an editing station with an analog video-recorder. The transmissions were recorded on analogous videotape as they reached the editing station. The station was controlled by the researcher and was located in a laboratory adjacent to the observational classroom but outside of the participants' view. Thus, the researcher was not physically present in the classroom during the observation sessions. The two recording videocameras were positioned just under the ceiling in the observational classroom, well outside of the participant's line of vision. The videocameras' position and noiseless functioning made them virtually inobtrusive. Care was taken to de-emphasize to the extent possible the fact that the classroom was being observed, both to ensure the participants' comfort and to preserve as much as possible the authenticity of their behavior and interactions.

The researcher observed classroom interaction as it was being recorded on videotape,
then reviewed the tape again and selected episodes of classroom interaction to be used as stimuli for the recall interviews. The selected episodes were digitized by means of video-editing software (PowerCut Pro for Mac) and transferred to a portable memory-disk. The following criteria were applied in the selection of the episodes: Each episode was a complete conversational exchange between a learner and the teacher, rather than a part of a longer exchange; each episode had a clearly marked beginning and end; and the quality of the audio-visual signal was good enough so as to not impede the learner’s understanding and recollection (Mackey et al., 2007). The selected episodes for each class comprised 16-27 episodes of differing length (4 to 28 seconds, M=18 sec), of which 90% were episodes of corrective feedback, and 10% were distractors – episodes of learner-teacher interaction unrelated to error correction, e.g. teacher-provided instructions or explanations, learner questions and comments, etc.

After the end of each observed class meeting, the learners were invited to participate in a stimulated recall interview. Only a few (5-7) learners from each class opted to participate in this stage of the study. The volunteers from each class were brought to a computer lab and assigned an individual computer station equipped with headphones, a microphone and audio-recording software (Hijack for Mac). Using a pull-down screen visible to all interviewees, the researcher projected the videotaped classroom episodes, taken from the learners’ class session earlier in the day. The episodes were projected one at a time, consecutively, in the order in which they happened in class. After each episode, the interviewees were asked to put on their headphones and record their recollection of the thoughts they'd had during the original classroom interaction. They were specifically instructed to record only the actual thoughts they had during the original interaction, to the extent that they remembered them, and whenever they found recollection impossible, to say they did not remember. As already discussed in Chapter II, the instructions are
a crucial element in obtaining valid and reliable data through stimulated recall interview (Gass & Mackey, 2000). To ensure that all interviewees were given identical instructions, the instructions were typed beforehand and read to the participants (see Appendix D for a copy). After 30 seconds, the interviewees were interrupted, if they had not already completed their recall, and a new episode was projected.

Each teacher was taken to a different laboratory for a stimulated recall interview that followed the same format.

The interview technique described in the preceding section is known in SLA as stimulated recall (Gass & Mackey, 2000). Stimulated recall interviews are retrospective interviews in which audio or video stimuli are employed to elicit post-factum qualitative comments about performance. Stimulated recall interviews are a widely used measure for noticing in research on interaction (e.g., Adams, 2003; Egi, 2007, 2010; Kim & Han, 2007; Mackey 2002; Mackey et al., 2000; Mackey et al., 2001; Mackey et al, 2004; Mackey et al., 2007; Nabei & Swain, 2002; Yoshida, 2008; 2010). In selecting the most appropriate technique for data elicitation in this study, other measures of perceptions, noticing, and awareness that were considered along with stimulated recall, included think-aloud protocols (Bowles, 2010), online and offline questionnaires (Mackey et al., 2001; McDonough & Mackey, 2006; 2008), and immediate recall protocols (Egi, 2004; Philp, 2003). These measures, however, were ruled out because they would interfere excessively with classroom communication (think-alouds, immediate protocols) or would provide limited qualitative comments (questionnaires).

As discussed in Chapter 2, specific measures need to be taken to ensure that the data obtained through stimulated recall interviews are valid. These measures relate to the prevention of memory decay (the progressive loss of memory as more time passes after the event that is
being recalled) and fabrication (the formation of false memories as a result of the time distance between the original event and the recall and the introduction of new information related to the original event) which undermine the reliability of the data.

Minimizing the time-lag between the initial experience and the stimulated recall interview is crucial for both the quantity and the quality of the data obtained through such measures, as well as for maintaining the core validity of the data as accurate recall of the learners' actual thought processes rather than of their post-factum reconstructions (Gorin & Stone, 2001). For this reason, a number of special measures were taken to reduce the time-lag between the original exposure to the classroom interaction episodes and the stimulated recall sessions to the minimum, within the practical limits imposed by learner schedules and logistical considerations. These measures included conducting all stimulated recall interviews within 3 to 6 hours of the original exposure - as soon as practically possible - as well as scheduling all interviews before the participants' next Arabic language class or study session took place, to prevent them from being exposed to new information and classroom interaction that could blur their memory of the original episodes and create false memories.

Data preparation for quantitative and qualitative coding

All videotaped classes, classroom feedback episodes, and learner and teacher interviews were transcribed by the researcher\(^5\). Two separate analyses were applied: quantitative and

---

\(^5\) The IRB-approved protocol for this study (Georgetown University IRB-C, Approval # 2006-513) stipulated that only the principal investigator (the researcher) and the responsible participant (the researcher’s supervisor) would have access to any data that allow the identification of a participant. Thus, any manipulation of the data that involved access to the original video- and audio-data, and to the real names of the participants in the study was performed by the researcher. Such manipulations included: collection and storage of informed consent forms, questionnaires, and video- and audio-data; assigning participant codes and pseudonyms; transcription of
qualitative. The coding scheme used for quantitative data analysis is illustrated in Figure 3 and described in the following section. The qualitative analysis is presented in detail at the end of this chapter.

![Figure 3. Coding categories for quantitative analysis.](image)

classroom episodes; transcription of stimulated recall interviews, and the first round of the coding sequence.
Quantitative Analysis of the Data

The data were coded in two steps. In the first round of data coding, the researcher compared all transcripts of classroom feedback episodes to the original videotaped episodes, identified the addressee of each episode, and marked the episode with the participant’s code. All learners’ stimulated recall interview transcripts were then compared to the classroom episodes and coded for direction of the feedback. When the learner providing the stimulated recall comment was also the learner to whom the corrective feedback had been addressed, that comment was coded as “direct” - i.e., the learner was commenting on feedback directed at herself. All other comments were coded as “indirect” – i.e., the feedback in the commented-upon episodes was directed at other learners in the classroom.

In the second round of data coding, the transcripts of all classroom episodes, teacher stimulated recall interviews, and learner stimulated recall interviews were coded separately. The coding categories used will be described below and illustrated with examples from the data set.

Classroom episodes

The classroom episodes analyzed in this study were transcribed communicative exchanges between the teacher and one learner, in which the learner produced a non-target like utterance, the teacher provided corrective feedback that addressed one or more linguistic problems in the learner’s utterance, and the learner either responded to the teacher’s feedback in

---

6 To protect the participants’ identity, real names were not used in this study. Upon agreeing to participate, each participant was assigned a code, containing basic information about their role (learner or teacher), gender, and order in the recruitment sequence (a number which showed which level, class and section the participant was recruited from). In all references to the participants throughout this text (in examples, discussion, quotes, etc.) these codes were replaced with pseudonyms for ease of reading.
some way, or continued without response\(^7\). This exchange sequence sometimes lasted more than three turns in the cases when the teacher, unsatisfied with the learner’s response, continued to address the same error in subsequent turns, pushing the learner to further modify their output.

Classroom episodes were coded for direction, type of feedback, and linguistic target of the feedback.

**Type of feedback**

When coding for types of feedback, each turn provided by the teacher was treated as a separate instance of corrective feedback, so that there could be more than one type of feedback in a single feedback episode.

*Explicit.* Feedback in which the teacher explicitly stated that the learner’s utterance was problematic and/or provided metalinguistic information (e.g., explanations about grammar or vocabulary, grammatical terminology) was coded as explicit. Example 6 illustrates a typical feedback episode containing explicit feedback.

Example 6: Explicit feedback, beginning classroom.

L: \( naHnu \ *nadxiin kathiir naHnu um \ *nadxiin? \)

   (we *smoke a lot…we …*smoke?)

   we *smoke a lot.. we *smoke?

T: \( nuu…? al-wazn ath-thaanii. \)  
   (Explicit correction – metalinguistic information)

   [we]…? Form two.

\(^7\) As previously stated, all types of response were treated equally. They were not analyzed as indications of learner understanding or lack thereof, because such an analysis is beyond the scope of this dissertation.
L:  nudaxxin

We smoke.

In this example, the learner (L) produced a non-targetlike 1st person plural form of the verb “to smoke” – *nadxiin instead of nudaxxin. The teacher (T) used grammatical terminology to point the learner to the correct morphological pattern for the needed verb, form II. This helped the learner produce the correct form, nudaxxin.

Negotiation. Feedback was coded as negotiation in the instances when the teacher indicated that they did not understand the learner’s utterance by either repeating the original non-targetlike utterance with interrogative intonation, or asking questions to ascertain what the learner’s intended meaning was.

Example 7: Negotiation, beginning classroom.

L:  *Saghiir ’uxt ’indii .. ’indii *Saghiir ’uxt … ’induki?

(*little-masc sister I-have.. I-have *little-masc sister… you-have?)

I have… *younger sister…I have…do you have?

T:  hal ’indii…?  (Negotiation – clarifying question)

do I have…?

L:  ’induki *Saghiira’ uxt?

(you-have little-fem *sister?)

do you have a * younger sister?

Declarative recast. Declarative recasts were instances of feedback in which the teacher recast the learner’s non-targetlike utterance, i.e., repeated the original utterance replacing the non-targetlike features with targetlike forms.
Example 8: Declarative recast, beginning classroom.

L:  

\textit{mundhu *waaHid sana}  

(Morphosyntactic error – word order and agreement)  

(since *one-masc. year)  

for *one year  

T:  

\textit{sana waaHida}  

(Declarative recast)  

for one-fem. year  

for one year

\textit{Interrogative recast.} Interrogative recasts were instances of feedback in which the teacher recast the learner’s non-targetlike utterance with interrogative intonation.

Example 9: Interrogative recast, beginning classroom.

L:  

\textit{huwa mutaxaSSiS fii ... fi-l.... *fi-l-fiziika?}  

(he-specializes in.. in-the.. in-the-*fizika?)  

he specializes in *fizika?  

T:  

\textit{fiiziyaa’?}  

Physics?

\textit{Combination.} Instances in which a single turn by the teacher contained more than one type of feedback addressing the same error, were coded as combination.

\textit{Linguistic target of the feedback}

\textit{Syntax.} The linguistic error which triggered the feedback was at the sentence level. Errors
in the data pool included word order, appropriate use of question words (e.g., hal), appropriate use of negation in the past and present tense (laysa, ma kaana).

Example 10:

L:  

*hunaaka *waaHid shay laazim nunaaqish  

(Syntactic error – word order)

(there is *one thing necessary we-discuss)

there is *one thing we should discuss

T:  

*shay’waaHid, shay’waaHid  

(Recast)

one thing, one thing

In Example 10 above, the learner used the incorrect word order for the phrase *shay’ waaHid (one thing). The teacher recasted the non-targetlike word order. This corrective feedback episode was coded as one with a syntactic target.

*Morphology.* The teacher’s feedback was directed at a linguistic error at the level of morphology: e.g., related to verb conjugation, appropriate choice of verb form, tense choice or formation, negation, subject-verb agreement, gender, number, definiteness, noun-adjective agreement, and the genitive construction for possession (*al-'Daafa*).

Example 11:  

(Morphological error – conjugation)

L:  

*hiya... hadhaa shay' taqliidii.. 'aDhunnu... wa laakin hiya lam tatakallam... wa... 'indama *takallamt

(she... this thing traditional I-think... and but she neg-she-spoke and when *I-spoke)

She – this is something traditional, I think, but she hadn't spoken and when *I spoke
In Example 11, the learner used a non-targetlike conjugation form for the past-tense verb *takallama* – 1st person singular *takallam* (I-spoke), rather than the 3rd person singular feminine form *takallamat* (she-spoke). Therefore, this feedback episode was coded as one with a morphological target.

*Lexis.* This category included feedback directed at the following: use of the verbal noun instead of the verb, incorrect form of the irregular plural, incorrect choice or omission of prepositions, use of words in any language other than Arabic, omission of words.

Example 12:

L:  

*fii bayt...fii bayt* *outside*  

in house... in house *outside*  

T:  

*xaarija l-jaamiʿa*  

outside the university  

L:  

*xaarija l-jaamiʿa, naʿam*  

outside the university, yes
In the above example, the learner did not know the Arabic equivalent of a lexical item, “outside”, and used the English word instead, producing a non-targetlike utterance. The teacher supplied the correct Arabic vocabulary item, xaarija (outside). This episode was coded as a lexical episode, because the target of the correction was a non-Arabic word.

**Phonology.** Feedback in this category was directed at a linguistic item pronounced in a non-targetlike way.

Example 13:

L: ʿindii Saff *kamsat ʾayyaam
(I have class *five days)
I have classes *five days [a week]

T: xamsat ʾayyaam, xamsat ʾayyaam
five days, five days

In this example, the teacher corrected the learner’s non-targetlike pronunciation of the word xamsa (five). Thus, the episode was coded as one with a phonological target.

**Teacher recall interviews**

The teachers’ recall comments were used to compare and verify the researcher’s coding
of the intended linguistic targets of the classroom feedback. Whenever there was a mismatch, the
teacher’s comment was considered a more valid indication of the intended target and used
instead of the researcher’s.

The teachers’ recall interview comments were coded for linguistic content (i.e., the type
of linguistic error that, in the teacher’s opinion, had triggered their feedback).

*Linguistic content*

*Syntax.* This category included comments which mentioned syntax as the
linguistic target of the feedback.

Example 14: Comment with syntactic focus, Teacher T4:

“Yah, very basic mistake, very common mistake, shouldn't be there – *waaHid shay*
[*thing-one], one of the things that I feel shouldn't be there at that level, you know, and
one of the things that I don't accept...”

*Morphology*

Example 15: Comment with morphological focus, Teacher T4:

“Again, basic conjugation stuff: *hunna yuridna, hunna yuriiduuna* [they-fem.-pl. want-
fem.-pl., they-fem.–pl. want-masc.-pl.], you know, *hunna yuridna* [they-fem. want-fem.-pl.],
again it's some of the basic items in conjugation, *fii aS-Sarf* [in the morphosyntax].”

*Lexis*

Example 16: Comment with a lexical focus, T4:

“That was just a singular-plural kind of mistake which is not common and because the
word *Hamaa* [mother-in-law] is not a common word, I mean, how many times do we use it, it's not like *kitaab-kutub* [book-books] or *jaami’a-jaami’aat* [university-universities], and I think most of the students didn't know the plural, so I said *Hamawaat* [mothers-in-law] and the student herself was at one point struggling to come up with the correct plural form and I want to give her that aid, so she can feel more comfortable in the conversation.”

**Phonology**

Example 17: Comment with a phonological focus, T3:

“I .. was trying to correct his pronunciation of the word ‘*iqtiSaad* [economy] ..”

**No content**

Example 18: No content comment from T2:

“I have no idea what we’re doing here.”

**Learner recall interviews**

Comments provided by the learners during their stimulated recall interviews underwent three stages of coding. First, as already discussed at the beginning of this section, they were coded for direction during the first round of general data coding, in which the learner to whom the feedback was directed was identified. Second, learner comments were coded for their linguistic focus. At the third and final coding stage, the learners’ comments were compared to the actual classroom episodes they referred to, and coded for awareness of correction and awareness
of target. The second and third coding stages will now be described and illustrated with examples from the data.

Comments that contained information about the linguistic focus of the feedback episodes were further categorized into several categories depending on the linguistic feature they mentioned. Comments that did not contain such information were also categorized according to the extent to which they constituted valid recall.

Linguistic content

Syntax. The comments in this category mentioned grammatical rules related to sentence structure and word order.

Example 19: Comment with syntactic focus, advanced learner:

“I don't remember what was happening, like, I don't even remember it happening, until she said **waaHid shay** [*thing one] and that just caught my attention before he even corrected it because it just... Dunno, it seems wrong and I hadn't even noticed until a couple of weeks ago how many people were saying, like, **waaHid rajul** [*man one], **waaHid shay** [*thing one], whatever, but then once Teacher D had corrected that one, so she sort of should've seen it...”

Morphology. Comments in this category were related to morphology and mentioned verb forms, particles, verb conjugation, tense, agreement, definiteness and **al- iDaafa**.

---

8 In this study, the data were elicited using a stimulated recall procedure. One caveat of using a retrospective procedure of this kind is that the data are subject to memory decay – i.e., the participants may be unable to recall the original thoughts that the researcher is trying to elicit. Whenever a participant stated that they did not remember their original thoughts, their comment was considered non-valid recall for the purposes of this study, since such a statement could not be interpreted to reflect in any way on the participant’s understanding of the content of the corrective episode. “No recall” comments were not used in the statistical analyses of the data.
Example 20: Comment with morphological focus, beginning learner:

“I remember after – I made the same, like, I'd made the same mistake, I said tadhhab [you-masc. go] instead of tadhhabiina [you-fem. go] and you know she's a girl, obviously [laughs] and I was like – after the teacher corrected me, I was like, oh yeah, shoot! - again...”

*Lexis.* Comments in this category focused on a lexical item and mentioned the meaning of a known or unfamiliar word, English translation of Arabic vocabulary items, or the inability to retrieve the needed vocabulary item.

Example 21: Comment with lexical focus, beginning learner:

“Everyone seems to mix up 'asDiqaar' [friends] and 'iqtiSaad [economy]. It happened like three times, I think.”

*Phonology.* Comments were categorized as related to phonology if they mentioned the pronunciation of a specific vocabulary item or commented on the pronunciation of Arabic in general or on the learner’s perception of their own pronunciation.

Example 22: Comment with phonological focus, beginning learner:

“Um, so that is me speaking and, uh, when professor corrected me I remember that there was, like, an 'ayn there, so I was, like, trying to remember how to pronounce it.”

*Unspecified*

Often learners mentioned in their comments that a correction had been made but did not specify the target of the correction. Such comments were coded as unspecified.
Example 23: Unspecified comment (no specific linguistic focus), advanced learner:

“I think I did the same mistake earlier that day as well in one of the one or two times that I spoke. Also, again on her comments I was like “huh?” - again not in agreement with them – but there wasn't enough time at the end of the hour and I didn't really feel like talking enough to contest any of this and to contribute to the discussion.”

Example 24: Unspecified comment, advanced learner:

“I wish Zoe would say the right thing, cause he keeps repeating the same stuff over and over again trying to get her to say, you know, what he wants her to say correctly, what she said incorrectly, so he wants her to correct herself and she wasn't picking him up and I was, like, thinking, oh my God, hurry up so we can move forward... and, um, I think that was it.”

No content:

Comments which contained no linguistic information were coded according to their validity as recall evidence.

Unrelated. Whenever the learner made a comment which neither contained linguistic information nor mentioned a correction, such comments were categorized as unrelated. They were considered valid evidence for recall and included in the analysis of the results.

Example 25: Unrelated comment, beginning learner:

“Um... I was actually trying to make up the question she was asking him in my head, in

---

9 “Unrelated” comments were considered valid recall comments for the purposes of the study, because they did contain information about the learner’s thoughts at the time of the original interaction – the fact, that these thoughts were not related to the content of the episode, does not invalidate these comments as data - it simply shows that learners could be thinking about matters, unrelated to the classroom interaction.
No recall. The absence of comment or comments in which the learner stated that they did not remember their thoughts at the time when the original interaction took place, were categorized as “no recall”. Since these comments were a clear indication of the lack of recall, they were considered, for the purposes of the current study, identical to missing data, and excluded from the analysis of the results (Grotjahn, 1987).

Example 26: No recall comment, advanced learner:

“I don't remember this part of class. Any part of this section, this clip. Sorry.”

Awareness

As a final step, the learners’ comments were compared to the actual classroom episodes in order to check for the presence or absence of awareness. Comments, in which the linguistic focus, mentioned by the learner, overlapped with the linguistic focus of the correction, identified by the teacher, were coded as demonstrating awareness of target. Comments, in which the learner mentioned explicitly that a correction had been made, even if they did not identify the target of the correction, were coded as demonstrating awareness of correction. Examples 27 and 28 illustrate the two categories:

Example 27. Awareness of target.

Classroom episode:

L:  *haadith *al-sayyaara  (Morphological error – definiteness)

(accident *the-car)

*The car accident
T: Keep it indefinite (Explicit correction)
L: Okay
T: “A” car accident, haadith sayyaara (Explicit correction)
L: haadith sayyaara... haadhihi sana?
   (accident car... this year?)
   A car accident .. this year?

Stimulated recall comments:
Teacher: “Um… with indefinite ‘iDaafas … they don’t like them. They got used to making definite ‘iDaafas, so... it was almost every student – every student, as I was walking around [to monitor the learners’ preparation before the activity], they had the same thing, they wanted to make a definite ‘iDaafa. So I just make sure that they keep it indefinite in there.”

Learner (recipient of feedback): “This was on the second question I was asked and I was trying to mentally note to keep the word as-sayyaara, the car, indefinite.”

In the above example, the teacher identified indefiniteness as the focus of the corrective intervention. The learner’s comment shows that he had understood the target of the feedback correctly – the second word in the ‘iDaafa, sayaara [car], had to stay indefinite. Therefore, this episode was coded as positive for awareness of target.
Example 28. *Awareness of correction.*

L: Okay, let’s see… um, Karl, *ma’a man *yaskunuun? (Morphological error –
(with-whom * they-live?)
with whom do *they live?

T: *taskun* (Recast)
(you-masc. live)
you live

L: *taskun* (you-masc. live)
you live

**Stimulated recall comments:**

Teacher: “When conjugating for “he” – no, for “you”, sometimes the students make the
mistake in conjugating for “he” and that happens very often.’

Learner (observer of the correction): ‘The teacher corrected Nathan a little bit and then he
was able to figure it out.’

In this example, the learner did not mention the linguistic focus of the correction but did
explicitly state that she had observed a correction. This episode was coded as positive for
awareness of correction.

**Inter-rater reliability**

The researcher coded 100% of the data. To assess the reliability of the coding categories,
a subset containing 36% of all data (31% of classroom episodes, 30% of learner stimulated recall
comments, and 70% of teacher stimulated recall comments) was coded by an external rater and
compared to the researcher’s coding. The external rater received the following training: (a) she was given information about classroom feedback, and examples of classroom feedback episodes illustrating the coding categories for target and type of feedback, used in the study; and (b) she was given information about the stimulated recall technique and the coding categories for stimulated recall comments used in the study. After the information session, the external rater practiced coding classroom feedback episodes and stimulated recall comments, which were not included in the analysis.

Familiarity with the content of the classroom feedback episodes might influence the external rater’s coding of the stimulated recall comments (Mackey et al., 2000). To avoid contamination of the coding, following Gass & Mackey (2000), the following measures were taken: (a) the stimulated recall data and the classroom episodes selected for coding did not come from the same class; and (b) the data were coded in the following order: (1) learner stimulated recall comments, (2) teacher stimulated recall comments, and (3) classroom feedback episodes.

For a more rigorous assessment of inter-rater agreement, both simple percentage and Cohen’s Kappa coefficient were used. Simple percentage indicated a very high degree of agreement between the two raters, with values ranging from 91% to 100%. Cohen’s Kappa coefficient values ranged from .824 to 1.000, indicating excellent inter-rater reliability (Cohen, 1992). The reliability scores are summarized in Table 3.
Table 3. Reliability of Coding

<table>
<thead>
<tr>
<th>Coding category</th>
<th>Simple % agreement</th>
<th>Cohen’s Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target of feedback</td>
<td>91</td>
<td>.824</td>
</tr>
<tr>
<td>Type of feedback</td>
<td>100</td>
<td>1.000</td>
</tr>
<tr>
<td>Learner comments – awareness of correction</td>
<td>93</td>
<td>.880</td>
</tr>
<tr>
<td>Learner comments – awareness of target</td>
<td>93</td>
<td>.880</td>
</tr>
<tr>
<td>Teacher comments</td>
<td>92</td>
<td>.883</td>
</tr>
</tbody>
</table>

Qualitative Analysis of the Data

The narrative recall data obtained from the learners were qualitatively analyzed by means of inductive thematic analysis techniques (Attride-Stirling, 2001; Rubin & Rubin, 1995; Taylor & Bogdan, 1984). The purpose of thematic analysis is "finding and marking the underlying ideas in the data, grouping similar information together, and relating different ideas and themes to one another" (Rubin & Rubin, 1995, p. 229). To achieve this, the analysis proceeded as follows: As a first step, the researcher read closely all learner recall transcripts, identified all ideas – or basic themes – related to the research question, and physically copied each theme onto a separate card. In the next step, the coded basic themes were compared, patterns of distinction or overlap were identified, and cards containing similar basic themes were grouped together. The unifying - or organizing - theme for each such group became this group's label. Finally, in the third step, all
labels were further analyzed through cross-category comparisons, and converted to more abstract constructs, or global themes. Figure 4 illustrates the resulting thematic hierarchy.

Figure 4. Qualitative analysis – organization of themes.

Only episodes in which the learner was an indirect participant were used for this analysis, because of the assumption that in episodes directed at the reporting participant it would be difficult to disentangle a learner's motivation to listen stemming from a feature of the episode, from their motivation to listen as a response to the discursive demands of the conversation in which they had the role of an interlocutor and were under an obligation to provide a response.

In other words, when a learner was the direct recipient of the feedback, the learner’s role as participant in the teacher-learner dialogue was assumed to be a sufficient explanation for their attentiveness to the corrective feedback episode. At the same time, a learner who was merely witnessing a corrective feedback episode occurring in the classroom, and was under no
expectation to participate directly in it, could have a number of reasons for consciously directing their attention to, or away from, the correction. It was these reasons that were the focus of the qualitative analysis described in this section.

This chapter described the research questions that guided the study as well as the study design, the data collection procedures, instruments, and coding schemas. The next chapter will present the quantitative and qualitative findings from the study.
CHAPTER 4: RESULTS

Quantitative Results

In this section, the data collected in the study are described and analyzed in relation to the research questions, using quantitative methods. Descriptive statistics, as well as results from the inferential statistical tests, are reported.

Since the variables in the study were categorical, a nonparametric test of significance (Pearson’s chi-square test for independence), was deemed most suitable for the testing of all hypotheses (see, for example, Hatch & Lazaraton, 1991). The Fisher exact test was supplemented for Pearson’s chi-square whenever appropriate (i.e., in the case of contingency tables with only one degree of freedom). In all tests exact probability levels were reported, rather than a pre-set alpha level (Wright, 1997). Whenever the analysis involved variables with more than two values, Haberman’s adjusted residuals (-2< or >2) were used to identify the sources of significance. Following Wilkinson & et al (1999), measures of the strength of association (Cramer’s V) were also reported whenever appropriate, as they can shed light on the practical meaning of statistically significant results (Brown, 1988; Hatch and Lazaraton, 1991).

All inferential statistical analyses were conducted on SPSS 15.0 for Windows.
Description of the data

The data collected in this study consisted of a total of 247 classroom feedback episodes. Of those, 84 episodes (34%) were selected on the basis of the criteria described in Chapter 3, Method: Data collection procedure, and used as stimuli for learner recall interviews. Thirty-one stimulated recall interviews were conducted, 26 with beginning learners, and 5 with advanced learners, yielding a combined total of 524 comments. Each interview included between 15 and 23 episodes of corrective feedback ($M = 16.9$, $SD = 3.07$). All comments were transcribed and coded using the coding schema described in 3.7. Data coding for quantitative analysis. After the coding, 443 (84.5%) of all comments were judged valid recall data, and included in the data set used for subsequent analyses. The remaining comments (15.5%, $n = 81$) contained a clear indication that recall had failed (e.g. “I don’t remember what I was thinking.”). In keeping with the methodological assumptions of the recall method (Simon & Ericsson, 1984), these comments were considered non-valid data for the purposes of this study, which addressed only the content of the participants’ comments but not the likelihood of the recall (Grotjahn, 1987). These non-valid data were excluded from the calculations in the inferential analysis but were kept in the data charts and figures with the label “No recall”. Thus, the final data set used for the inferential analysis was comprised of 84 recall feedback episodes, and 443 recall comments ($M = 14.29$, $SD = 3.23$). Sixty-eight of the recall episodes came from the four beginning classrooms, and 16 episodes came from the advanced classroom. The 84 recall episodes will now be described.
Recall data set for inferential analyses

Table 4. Types of feedback in recalled corrective feedback episodes

<table>
<thead>
<tr>
<th>Type of feedback</th>
<th>Beginner classrooms</th>
<th>Advanced classroom</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Explicit feedback</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Elicitation</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Declarative recast</td>
<td>44</td>
<td>64</td>
</tr>
<tr>
<td>Interrogative recast</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Negotiation</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Combination</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>68</td>
<td>100</td>
</tr>
</tbody>
</table>

Note. n = raw frequency. Dashes indicate the absence of feedback with these characteristics from the sample.

As can be seen in Table 4 and Figure 5, the dominant feedback move in the recall episodes, at both proficiency levels, was the declarative recast (64% of the time in the beginning classrooms, and 75% of all corrective feedback in the advanced classroom). This distribution is consistent with findings from a large body of existing research that identify the declarative recast as the most frequent type of classroom corrective feedback (Lyster & Ranta, 1997; Panova & Lyster, 2002; Seedhouse, 1999). Combination followed in frequency at the rate of 14% at the beginning level and 19% at the advanced level. Interrogative recast, explicit feedback, elicitation and negotiation occurred rarely; together they accounted for 22% of the feedback moves in the
beginning classrooms. The observational samples from the advanced classroom contained no examples of elicitation, negotiation, or interrogative recast episodes, and only a small amount of explicit feedback episodes (6%).

![Types of feedback - Beginner classrooms](image)

![Types of feedback - Advanced classroom](image)

*Figure 5. Distribution of feedback by type*

The distribution of linguistic targets in the recall episodes is shown in Table 5 and Figure 6. At both proficiency levels, over half of the recall episodes targeted morphological errors (59% for the beginning classrooms, 62% for the advanced classroom). The feedback episodes with lexical targets followed in frequency - 20% in the beginning classrooms and 25% in the
advanced classroom. Phonological errors were the targets of feedback less often, at the rate of 14% in the beginning classrooms and only 6% in the advanced classroom. Finally, syntactic errors were rarely the focus of corrective feedback in the observational sample (7% for the beginning classroom, 6% for the advanced classroom).

Table 5. *Linguistic targets of recalled feedback episodes*

<table>
<thead>
<tr>
<th>Target of feedback</th>
<th>Beginner classrooms</th>
<th>Advanced classroom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax</td>
<td>5 7</td>
<td>1 6</td>
</tr>
<tr>
<td>Morphology</td>
<td>39 59</td>
<td>10 62</td>
</tr>
<tr>
<td>Lexis</td>
<td>14 20</td>
<td>4 25</td>
</tr>
<tr>
<td>Phonology</td>
<td>10 14</td>
<td>1 6</td>
</tr>
<tr>
<td>Total</td>
<td>68 100</td>
<td>16 100</td>
</tr>
</tbody>
</table>

*Note. n = raw frequency.*

No attempt had been made to control the type and linguistic target of the recall episodes, with the aim to preserve the authenticity of their distribution in each specific classroom interaction. A comparison between the type and target distributions in the selected recall episodes as well as in the complete data set showed that the distributions of linguistic targets and types of feedback in the recall episodes mirror those of the larger sample. Therefore, the selected recall episodes were considered a representative sample of the classroom feedback episodes observed in this study.
Research Question 1

The first research question asked if the learners were a) aware of the corrective feedback, occurring in their classroom during oral interactive activities, and, b) aware of the linguistic targets of this corrective feedback. Awareness of correction was operationalized as an explicit
statement in the learner’s recall report that a mistake and/or correction had been made. As Table 6 shows, 54.4% of the comments from beginning learners and 78.8% of the comments from advanced learners, contained indication of awareness.

Table 6. Awareness of corrective intent by learner’s proficiency level

<table>
<thead>
<tr>
<th>Recall comments</th>
<th>Proficiency level</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beginning</td>
<td>Advanced</td>
</tr>
<tr>
<td>+ Awareness</td>
<td>205</td>
<td>54.4</td>
</tr>
<tr>
<td>- Awareness</td>
<td>172</td>
<td>45.6</td>
</tr>
<tr>
<td>Total</td>
<td>377</td>
<td>100</td>
</tr>
</tbody>
</table>

Note. n = raw frequency.

As the numbers in Table 6 show, both beginning and advanced learners reported being aware of a correction more than half the time. One-way Pearson chi-square statistical tests showed that the number of the comments indicating awareness for each group was significantly larger than the number of comments which could not be interpreted to show awareness of correction (for beginners $\chi^2 (1, N = 377) = 2.889$, df (1), p=.089; for advanced $\chi^2 (1, N = 66 ) = 21.879$, df (1), p=.000.

Awareness of linguistic target was operationalized as overlap between the linguistic focus of the learner’s comment, and the linguistic target of the corrective feedback episode. As shown in Table 7, advanced learners indicated awareness of the linguistic target of the feedback 40.9%
of the time, while beginning learners were aware of the target of the feedback only 26.7% of the
time (see Table 6). This difference was also significant both for the beginning learners $\chi^2 (1, N = 377 ) = 79.87$, df(1), p=.000 and for advanced learners $\chi^2 (1, N = 66) = 2.182$, df(1), p=.140.

Table 7. Awareness of linguistic target by learner’s proficiency level

<table>
<thead>
<tr>
<th>Recall comments</th>
<th>Proficiency level</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beginning</td>
<td>Advanced</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>+ Awareness</td>
<td>102</td>
<td>27</td>
<td>129</td>
<td>29</td>
</tr>
<tr>
<td>- Awareness</td>
<td>275</td>
<td>39</td>
<td>314</td>
<td>71</td>
</tr>
<tr>
<td>Total</td>
<td>377</td>
<td>66</td>
<td>443</td>
<td>100</td>
</tr>
</tbody>
</table>

Note. n = raw frequency. Numbers in parentheses represent percentage of the total number of recalled episodes.

These results confirmed the hypothesis that learners in both groups will show relatively high rates of awareness of correction, following Mackey (2002) and Lyster & Mori (2006). It is interesting to note the reported rates of awareness by both groups. The awareness rate of advanced learners approached the very high numbers (70-90%), reported in laboratory conditions (Mackey, 2002; Philp, 2003), suggesting that learners at this proficiency level were less susceptible to interference from contextual distractors.

Learners were also more likely to report awareness of the correction than of the target. There are two ways to account for this finding: First, the difference may be due to limitations of the data collection technique – the study relied on retrospective data, which reflects the rate of
reported awareness but not necessarily the rate of actual awareness. Second, the discrepancy may be due to the level of expected cognitive engagement. For a learner in a feedback-heavy environment who is primed to interpret the teacher’s contribution as corrective, identifying the feedback as a conversational move may be easy – second nature. Identifying the target of the feedback, however, takes some level of knowledge about linguistic structures, and the activation and application of that knowledge in addition to recognizing the feedback move, and so it requires a greater cognitive effort.

**Research Question 2**

The second research question asked if there was a relationship between the proficiency level of the learners, and their awareness of a) the corrective intent of the feedback, and b) the linguistic target of the feedback.

Advanced learners indicated awareness of the corrective intent of the feedback 78.8% of the time. For beginning learners, the percentage of reported awareness was much lower: 54.4% of the time (see Table 8). The results from a Pearson chi-square test indicated that the difference was significant ($\chi^2 (1, N = 443) = 13.921, \ p = .000$).
The table shows the awareness of corrective intent by learner's proficiency level. The data is presented in a tabular format with columns for recall comments, proficiency level, and percentage. The table includes rows for + Awareness, - Awareness, and Total. The percentages for each group are calculated based on the total recall episodes.

### Table 8. Awareness of corrective intent by learner's proficiency level

<table>
<thead>
<tr>
<th>Recall comments</th>
<th>Proficiency level</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beginning</td>
<td>Advanced</td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ Awareness</td>
<td>205 54.4%</td>
<td>52 78.8%</td>
<td>257 58%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Awareness</td>
<td>172 45.6%</td>
<td>14 21.2%</td>
<td>186 42%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>377 100%</td>
<td>66 100%</td>
<td>443 100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** n = raw frequency.

Advanced learners indicated awareness of the linguistic target of the feedback 40.9% of the time, while beginning learners were aware of the target of the feedback only 26.7% of the time (see Table 9). This difference was also significant ($\chi^2 (1, N = 443) = 5.525$, Fisher’s Exact Test p = .026).

### Table 9. Awareness of linguistic target by learner’s proficiency level

<table>
<thead>
<tr>
<th>Recall comments</th>
<th>Proficiency level</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beginning</td>
<td>Advanced</td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ Awareness</td>
<td>102 26.7%</td>
<td>27 40.9%</td>
<td>129 29%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Awareness</td>
<td>275 73.3%</td>
<td>39 59.1%</td>
<td>314 71%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>377 100%</td>
<td>66 100%</td>
<td>443 100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** n = raw frequency. Numbers in parentheses represent percentage of the total number of recalled episodes.
These results confirm the proposed hypothesis that advanced learners will report higher rates of awareness for both correction and target.

Cramer’s V was used to measure the strength of association between proficiency level and accurate perceptions about corrective intent ($V = -0.176, p = .000$), and target ($V = -0.109, p = .026$). In both cases, the value of this measure was interpreted as low to lower than typical (Cohen, 1992).

Cramer's V is a measure of percentage of variance accounted for - that is, it shows how much of the difference between the two groups can be attributed to the variable under examination. In statistical terms, an effect size of 10-20% is a rather small one. A theoretical model that explains only a tenth or a fifth of the observed differences would be considered a weak model of low explanatory power, because it leaves unexplained a very large proportion of differences between the groups that can be due to the influence of any number of factors. The current study, however, did not aim to arrive at a sound theoretical model for the variance observed in the beginning and advanced classrooms. Such an aim could be achieved only by an experimental study in which all variables have been identified on the basis of existing empirical research, and carefully controlled. Because feedback in the Arabic foreign language classroom has not received much attention in previous research, such a controlled study could not be carried out for lack of validating empirical base. Instead, this study followed an explorative descriptive design with the goal to identify possible factors that influence learners' awareness of corrective feedback. In this sense, the theoretical model employed in the study was an open one, based on the assumption that interactions in a classroom environment are potentially influenced by a very large number of factors related to the setting, the communicative situation, and the participants in
the interaction. When we view the study in this light, it becomes obvious that interpretations based on the practical significance of the results would be more meaningful and informative than a focus on theoretical interpretations. In practical terms, a variable whose influence explains 10 to 20\% of the variation between the groups - out of a much larger pool of potential variables in this uncontrolled setting - is an important finding.

Consequently, while in this case the statistical interpretation of the value of Cramer’s V = .148 was that it showed a weak association between the linguistic target of the episode and the learners’ awareness of the correction, its suggested practical interpretation is that this was a value of high significance.

Research Question 3

The third research question asked if the learners’ awareness of the corrective intent of feedback was related to characteristics of the feedback. The following feedback characteristics were the focus of this study: a) the type of feedback, b) the linguistic target of the feedback, and c) the direction of the feedback in relation to the reporting participant.

As Table 10 and figure 7 show, at the beginning level all types of corrective feedback yielded similar rates of reported awareness – between 50 and 68\%, with the exception of Negotiation, which was noticed 82\% of the time.
This is a surprising finding, since Negotiation is a type of implicit corrective feedback which tends to be overlooked by language learners. In an interesting contrast, another type of implicit feedback – the declarative recast – showed the lowest rates of awareness (50%) of all corrective feedback types. Such low noticeability is consistent with theoretical and empirical research on learner awareness of recast (e.g., Lyster & Ranta, 1997). It may be the case that not all types of implicit feedback are created equal and that even feedback of typically low salience may become highly noticeable due to contextual factors.
### Table 10. Beginning and advanced learners’ awareness of correction by type of feedback

<table>
<thead>
<tr>
<th>Feedback type</th>
<th>Beginners + Awareness</th>
<th>Advanced + Awareness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit</td>
<td>15 (68%)</td>
<td>3 (60%)</td>
</tr>
<tr>
<td>Elicitation</td>
<td>14 (56%)</td>
<td>_</td>
</tr>
<tr>
<td>Declarative recast</td>
<td>120 (50%)</td>
<td>37 (77%)</td>
</tr>
<tr>
<td>Interrogative recast</td>
<td>20 (62%)</td>
<td>_</td>
</tr>
<tr>
<td>Negotiation</td>
<td>9 (82%)</td>
<td>_</td>
</tr>
<tr>
<td>Combination</td>
<td>27 (57%)</td>
<td>12 (92%)</td>
</tr>
</tbody>
</table>

*Note. Values represent raw numbers and percentage of recall comments. Dashes indicate absence of feedback with this characteristic from the data.*

The same unorthodox distribution of reported awareness can be seen in the data from the advanced classroom (Figure 8): There, the declarative recast – the only implicit feedback type in this data pool – yielded higher rates of awareness than explicit feedback (77% vs 60%, respectively).

105
Advanced learners’ awareness of correction by type of feedback

![Advanced learners' awareness of correction by type of feedback](image)

Figure 8. Advanced learners’ awareness of correction by type of feedback.

A chi-square test did not confirm that these differences are statistically significant ($\chi^2 (5, N = 377) = 7.938, p = .160$ for the beginners, and $\chi^2 (2, N = 66) = 2.561, p = .278$ for advanced learners) – a result that may have been due to differing numbers of corrective feedback episodes of each type. However, the trend suggested by the frequency counts of awareness at each level remains an interesting finding.

Awareness of linguistic target was operationalized as overlap between the linguistic focus of the learner’s comment, and the linguistic target of the corrective feedback episode. While some linguistic targets seemed more readily recognizable at the advanced level (e.g., Phonology – 100%, Syntax – 80%), a chi-square test showed that these differences were not statistically significant ($\chi^2 (3, N = 66) = 4.769, p = .190$).
At the beginning level, the rates of reported awareness of correction were similar across the four linguistic targets – just over half of the recall comments indicated awareness of correction, regardless of its target ($\chi^2 (3, N = 377) = 5.258$, $p= .262$). There was a slight advantage to corrective feedback episodes targeting lexical errors – those tended to be noticed 15% more than other lexical targets. Such a finding is consistent with previous research which suggests that lexical targets in the input tend to be noticed more frequently, possibly due to higher perceptual salience (Mackey, Gass & McDonough, 2000). However, as the chi-square statistic shows, this advantage was not statistically significant. Table 11 and Figure 10 present these results visually.
Table 11. *Beginning and advanced learners’ awareness of correction by linguistic target of feedback*

<table>
<thead>
<tr>
<th>Target</th>
<th>Beginners</th>
<th></th>
<th>Advanced</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ Awareness</td>
<td>- Awareness</td>
<td>+ Awareness</td>
<td>- Awareness</td>
</tr>
<tr>
<td>Phonology</td>
<td>29 (53%)</td>
<td>26 (47%)</td>
<td>4 (100%)</td>
<td>0</td>
</tr>
<tr>
<td>Morphology</td>
<td>107 (51%)</td>
<td>103 (49%)</td>
<td>35 (83%)</td>
<td>7 (17%)</td>
</tr>
<tr>
<td>Lexis</td>
<td>52 (65%)</td>
<td>28 (35%)</td>
<td>9 (60%)</td>
<td>6 (40%)</td>
</tr>
<tr>
<td>Syntax</td>
<td>13 (50%)</td>
<td>13 (50%)</td>
<td>4 (80%)</td>
<td>1 (20%)</td>
</tr>
</tbody>
</table>

*Note.* Values represent raw numbers and percentage of recall comments.

Fig. 10. *Beginning learners’ awareness of correction by target of feedback.*

Direction of the feedback, i.e., the commenting learner’s direct or indirect involvement in the corrective feedback episode, was significantly associated with awareness of correction, when the data were analyzed as a whole. A breakdown by proficiency level, however, shows a
different picture. While beginning learners were significantly more likely to report awareness of correction when the feedback was directed at them (82%), rather than at someone else in the class (52%), $\chi^2 (1, N = 377) = 9.400, p = .002$, Cramer’s $V = .158$ (low to lower than typical), for advanced learners the direction of the feedback clearly made no difference. They were equally likely to notice the correction in a direct episode (80%) and in an indirect episode (79%), $\chi^2 (1, N = 66) = .010$, Fisher exact test $p = 1.000$. Table 12 and Figure 11 illustrate the result.

Table 12. Beginning and advanced learners’ awareness of correction by direction of feedback

<table>
<thead>
<tr>
<th>Direction</th>
<th>Beginners</th>
<th></th>
<th></th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ Awareness</td>
<td>- Awareness</td>
<td>+ Awareness</td>
<td>- Awareness</td>
</tr>
<tr>
<td>Direct</td>
<td>23 (82%)</td>
<td>5 (18%)</td>
<td>8 (80%)</td>
<td>2 (20%)</td>
</tr>
<tr>
<td>Indirect</td>
<td>182 (52%)</td>
<td>167 (48%)</td>
<td>44 (79%)</td>
<td>12 (21%)</td>
</tr>
</tbody>
</table>

*Note. Values represent raw numbers and percentage of recall comments.*
Beginning learners’ awareness of correction by direction of feedback

Advanced learners’ awareness of correction by direction of feedback

Figure 11. Beginning and advanced learners’ awareness of correction by direction of feedback.

Research Question 4

The last research question addressed the relationship between the learner’s awareness of the linguistic target of the feedback, and a) the type of the feedback, b) the linguistic target of the feedback, and c) the direction of the feedback.
Beginning learners were significantly more likely to report awareness of the linguistic target in corrective feedback episodes involving Explicit feedback (46%) and Combination (43%) and less likely to do so in episodes, involving Declarative recast (22%). Therefore, in the case of beginning learners, this relationship was statistically significant $\chi^2 (5, N = 377) = 15.06, p = .010$, Cramer’s V = .200 (low to lower than typical). Table 13 and Figure 12 illustrate the results.

Table 13. Beginning and advanced learners’ awareness of target by type of feedback

<table>
<thead>
<tr>
<th>Feedback type</th>
<th>Beginners</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ Awareness</td>
<td></td>
</tr>
<tr>
<td>Explicit</td>
<td>10* (46%)</td>
<td>1 (20%)</td>
</tr>
<tr>
<td>Elicitation</td>
<td>5 (20%)</td>
<td>_</td>
</tr>
<tr>
<td>Declarative recast</td>
<td>53* (22%)</td>
<td>20 (42%)</td>
</tr>
<tr>
<td>Interrogative recast</td>
<td>9 (28%)</td>
<td>_</td>
</tr>
<tr>
<td>Negotiation</td>
<td>5 (46%)</td>
<td>_</td>
</tr>
<tr>
<td>Combination</td>
<td>20* (43%)</td>
<td>6 (46%)</td>
</tr>
</tbody>
</table>

*Note. Values represent raw numbers and percentage of recall comments. Dashes indicate absence of feedback with this characteristic from the data. Cells marked with asterisks contributed to a significant difference (Haberman’s adjusted residuals).

This finding is consistent with previous research, according to which more explicit types of corrective feedback are more successful in drawing the learners’ attention to features in the input, and implicit types of corrective feedback, such as recasts, are less likely to be noticed.
Advanced learners seemed to recognize most readily Phonological targets (75%) than Lexis (47%), Morphology (6%) or Syntax (40%), but a chi-square test revealed that the differences were not statistically significant ($\chi^2 (3, N = 66) = 2.599, p = .458$). In contrast, beginning learners were least likely to recognize phonological targets (13%), and most frequently reported awareness of lexical targets (39%). At this proficiency level, the relationship between the linguistic target of the feedback, and the learners awareness of this target, was
significant ($\chi^2 (3, N = 377) = 14.661, p= .005, \text{Cramer's } V = .197$ – low to lower than typical).

Table 14 and Figure 13 present these results visually.

**Table 14. Beginning and advanced learners’ awareness of target by linguistic target of feedback**

<table>
<thead>
<tr>
<th>Target</th>
<th>Beginners</th>
<th></th>
<th></th>
<th>Advanced</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ Awareness</td>
<td>- Awareness</td>
<td>+ Awareness</td>
<td>- Awareness</td>
<td>+ Awareness</td>
<td>- Awareness</td>
</tr>
<tr>
<td>Phonology</td>
<td>7* (13%)</td>
<td>48* (87%)</td>
<td>3 (75%)</td>
<td>1 (25%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphology</td>
<td>57 (27%)</td>
<td>153 (73%)</td>
<td>15 (36%)</td>
<td>27 (64%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexis</td>
<td>31* (39%)</td>
<td>49* (61%)</td>
<td>7 (47%)</td>
<td>8 (53%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntax</td>
<td>4 (15%)</td>
<td>22 (85%)</td>
<td>2 (40%)</td>
<td>3 (60%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. Values represent raw numbers and percentage of recall comments. Cells marked with asterisks contributed to a significant difference (Haberman’s adjusted residuals).*
Finally, analysis of the data showed a significant association between the direction of the feedback episode and the learners’ awareness of the linguistic target of the feedback. at the beginning level ($\chi^2 (1, N = 377) = 5.752, p = .016$, Cramer’s $V = .124$, lower than typical) but not for the advanced learners $\chi^2 (1, N = 66) = 2.80, p = .077$). Table 15 and Figure 14 present these results.

Table 15. 

| Direction | Beginners | | | Advanced | | | |
|-----------|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|
|           | + Awareness | - Awareness | + Awareness | - Awareness | + Awareness | - Awareness |
| Direct    | 13 (46%)   | 15 (54%)     | 7 (64%)      | 3 (36%)       |                |                |
| Indirect  | 89 (26%)   | 260 (74%)    | 20 (70%)     | 36 (30%)      |                |                |

*Note. Values represent raw numbers and percentage of recall comments*
Advanced learners’ awareness of target by direction of feedback

Beginning learners’ awareness of target by direction of feedback

Figure 14. Beginning and advanced learners’ awareness of target by direction of feedback.

Qualitative findings

In addition to the quantitative analysis already described in the previous sections, the narrative data were qualitatively analyzed using a thematic networks technique (Attride-Sterling, 2001). The three-step analysis proceeded as follows: First, the researcher read closely all transcripts and identified basic ideas, related to the research questions. Specifically, the ideas that
were the focus of this analysis were learner-reported reasons to pay attention to the observed corrective feedback episode, or to ignore it. Second, the basic ideas were compared and organized on the basis of similarities, and thematic labels were assigned to the organizing themes, around which cluster of similar ideas had formed. In the third and final step, cross-category comparisons were used to enable macro-, or global themes to emerge from the clusters of organizing themes. These global themes, defined as “super-ordinate themes that encompass the principal metaphors in the data as a whole” (Attride-Sterling, 2001, p. 388), constituted the core of each thematic network.

Only episodes in which the learner was a direct participant were used for this analysis, because of the assumption that in indirect-participation episodes it would be difficult to disentangle a learner's motivation to listen, stemming from a feature of the episode, from their motivation to listen as a response to the discursive demands of the conversation in which they had the role of an interlocutor and were under an obligation to provide a response. In other words, when a learner was directly involved in the feedback episode, i.e., he or she was the recipient of the feedback, the learner’s role as participant in the teacher-learner dialogue was assumed to be a sufficient explanation for their attentiveness to the corrective feedback episode. At the same time, a learner who was merely witnessing a corrective feedback episode occurring in the classroom, and was under no expectation to participate actively in it, could have a number of reasons for consciously directing their attention to, or away from, the correction. These reasons were the focus of the qualitative analysis, described in this section.
The three-step thematic analysis resulted in a nexus containing five thematic networks, organized around five global themes: Learner, Feedback Episode, Linguistic Information, Message, and External Distractors (Figure 15).

*Figure 15. Global themes identified in the thematic analysis.*

*Learner:* These comments referred to the interviewee’s attitude towards the participant who was the recipient of the corrective feedback in the observed episode. The major sub-themes included both positive and negative attitudes towards the language skills, knowledge, diligence, and personality of the learner in the center of the corrective feedback episode.

*Feedback episode:* These comments referred to the interviewee’s attitude – both positive and negative, towards the teacher's use of specific error-correction strategies.
Linguistic information: These comments referred to the linguistic information that was contained in the corrective feedback episode, particularly information related to the non-targetlike utterance which triggered the correction.

Message: These comments referred to the meaning of what the participant at the center of the corrective feedback episode was saying, or the informational content of their contribution.

External distractors: These comments referred to factors that were not part of the corrective feedback episode, or directly related to it, yet they were perceived by the interviewee to affect the quality of their engagement with the corrective feedback.

Each thematic network will now be described and illustrated with examples from the narrative data. Figure 13 illustrates the hierarchical organization of basic, subordinate, and global theme in the thematic network “Learner”.

Figure 16. Thematic network “Learner”.

118
Global theme “Learner”

Comments, related to this theme were the most common in the recall data for the study. In all, twenty out of the twenty-six interviewed learners contributed a combined total of 41 recall comments that discussed their positive or negative attitudes to the learner at the center of the feedback episode. The interviewee’s comments were organized in five subordinate themes, as follows:

“This learner is my friend.”

A small number of comments referred to an existing personal relationship with the learner at the center of the feedback episode, and posited this relationship as a reason for special attention to what the learner was saying, or empathy in cases where the learner was having trouble.

Example 27: “I do remember I was giving much more attention to the way he was saying things because he is a friend of mine and that I wanted to try to understand him, whereas I don’t have any incentive with the others – except for Zoe who’s the other one who is also .. who is also an acquaintance of mine who speaks very well. I remember that I really wanted to pay attention, so if there was something that I could give to the conversation, almost to show my support and interest in what he was saying. I was paying attention even though it was very difficult to grasp what he was saying.”
“This learner is good.”

A very common reason for paying special attention to the content and especially the form of the learner’s contribution, was the interviewee’s perception of the learner as particularly good at Arabic, with excellent language skills and expertise knowledge that the interviewee trusted and tried to emulate, and whose speech was a pleasure to listen to.

Example 28: “He probably sounds the best in Arabic – he’s a very good speaker, I love hearing him talk.”

Example 29: “She’s pretty good and she’s trying to say something complicated but she wasn’t sure how.”

Example 30: I was impressed that she knew how to say the past tense, because I had no idea how. And her accent is really good!

“This learner works hard.”

This topic was related to the previous, yet independent from it in that the interviewee’s comments focused not on the learner’s expertise or skills but on the perception of his or her diligence and hard work, which were typically viewed in a positive light. Interviewees stated frequently that they “felt really bad” for a learner who was having trouble expressing him- or herself, empathizing with their perceived difficulty or embarrassment.

Example 31: “I remember thinking – she knows what she has to say but she thinks too much, she just tries too hard to pronounce it correctly.”

Example 32: “She works hard, so I’m really disappointed that she didn’t do well.”
“This learner knows a lot, so I was sympathetic.”

Sometimes interviewees commented specifically on their empathy for a good learner with known expertise who happened to be having a hard time producing a correct utterance or expressing a thought well during the observed episode. Unlike the comments in the previous two subordinate themes, these comments put the focus on the interviewee’s sympathy, rather than the learner’s extent of knowledge or diligence.

Example 33: “I was thinking about the girl – she was nervous and although she knows everything, she’s tripping over herself.”

Example 34: “I felt really bad for him [for making a mistake]. Actually, he’s one of the best people in our class and I just kinda feel sort of bad for him sometimes.”

“I dislike this learner and/or their contribution.”

A number of comments referred specifically to negative feelings towards the learner who was at the center of the feedback episode. These feelings seemed to be unrelated to the particular situation or class but rather extended to a general animosity towards the learner and distaste for their personality or other characteristics, particularly their voice or intonation, rather than their language expertise.

Example 35: “That kid who’s speaking – I find him really obnoxious and annoying. So every time he opens his mouth, I just want him to shut up. It’s awful but it’s true. So when he speaks, I think about how I wish he wasn’t speaking.”
Example 36: “I wasn’t even trying to understand, cause whenever Paul says anything, I just try to pretend he doesn’t exist.”

Example 37: “Getting frustrated everyday like I do with Dina, cause her voice is like a lullaby. I’m just making fun of her in my head. I remember actually trying to follow what she was saying and then kind of… stopping. Just tired of listening to her.”

*Figure 17.* Thematic network “Feedback Episode”

Global theme “Feedback Episode”

As can be seen in Figure X, this theme was a combination of two opposite subordinate themes, which described the interviewee’s positive or negative feelings towards the teacher’s correction or overall behavior in the observed episode. An interesting feature of the two
subordinate themes is that the positive and negative attitudes seem to be closely related to the interviewee’s proficiency level. Out of the total of seven learners whose comments referred to the teacher, three reported a negative attitude – all of them advanced learners, and four had a positive impression of the teacher’s intervention – the majority of them, again three, were beginning learners.

“I disliked the teacher's correction.”

Advanced learners saw the classroom discussion as a conversation and were interested in the content of what their peers were saying just as much as – or perhaps more than – the grammatical correctness of their production. The teacher’s intervention was sometimes seen as an unnecessary interruption to the flow of the conversation.

Example 38: “I was annoyed at the teacher, because Melanie was trying to talk, trying to get something across, and he sort of stopped her in the middle for a grammar lesson. And I felt that sort of disrupted the flow of what she was trying to say. I was just like, “Oh, God! Why did he have to do that?! She was getting somewhere, we could understand her, why did he have to stop?!”

“I liked the teacher's correction.”

Beginning learners, on the other hand, were more likely to express positive feelings about the teacher’s correction. They tended to see the teacher’s intervention as a reassuring form of assistance.
Example 39: “Good job of correcting without really taking out of the conversation too much.”

Example 40: “I liked how she corrected him, because it reassured me with what I’ve been saying.”

Figure 18. Thematic network “Linguistic information”.

Global theme “Linguistic information”

The interviewees frequently commented on the linguistic content of the observed corrective feedback episode, and on their reasons to pay particular attention to specific linguistic features, which were not necessarily the features targeted by the corrective feedback but were, nevertheless, perceived by the learner as important knowledge that was either missing from their repertoire, or which they had been misinterpreting. This was the largest theme to emerge from
the recall data, with twenty-two learners contributing a total of 45 comments. Figure X presents visually the thematic network for this global theme.

“I have trouble with the same thing.”

Learners frequently mentioned that they felt motivated to pay more attention to corrective episodes focusing on linguistic features they themselves had not yet mastered. Hearing someone else’s error corrected was a way for them to learn in a non-face threatening way, without the anxiety of being the recipient of a correction themselves.

Example 41: “That was a good review for me, because I’m like, oh, wait, I’m making the same mistake! – which was the same mistake that he made.”

Example 42: “The girl here has trouble with conjugations, which I have trouble with, too, so as she was trying to figure it out, I was trying to figure it out with her.”

“I had just made the same mistake.”

In a related sub-theme, learners commented on their motivation to pay specific attention to an error they themselves had made during the same class period or activity, as a way of rehearsing and consolidating the new linguistic information they had received.

Example 43: “I made the same mistake earlier, so when she got corrected as well, I realized I hadn’t noticed that she used the possessive, and I hadn’t noticed when I used it.”
“This mistake had come up many times.”

Frequent errors tended to attract more attention and, at the same time, were more likely to be recalled in the interviewee’s reports. It seemed that the sheer repetition of the same linguistic information made these errors more readily noticeable in the stream of classroom speech.

Example 44: “That’s something a lot of us do a lot, so I heard her when she slipped the first time.”

“I was thinking of the correct way to say this.”

Often the interviewees commented that they were not paying attention to the corrective feedback episode at hand, because they were too concentrated on coming up with a hypothesis of their own. Conversely, interviewee’s comments sometimes indicated that when they had already formulated their own hypothesis on the correct way to say something, their attention to the classroom interaction and the feedback their peer was receiving, increased, as they listened for a confirmation to their hypothesis.

Example 45: “I remember thinking – I wonder if she’ll say it? – because it had been corrected, like, three or four times already. I remember anticipating the preposition and seeing if she’ll get it right.”
“This was new linguistic information, so I took note.”

New linguistic information attracted more attention as well, particularly if it was relevant to the activity in which the interviewee was involved, so they could easily see a practical application for it.

Example 46: “I also didn’t know which negation to use, so the teacher corrected her, so I wrote that down, cause I didn’t know that either.”

“I was thinking about how I would answer that question.”

Interviewees, both beginning and advanced, frequently reported rehearsing their own answers to the questions their peers were responding to.

Example 47: “While she was trying to get the correct dual form, I was trying to do the same in my head.”

“I was getting impatient, because this learner would not give the correct answer.”

Finally, when the interviewee did not find anything in the observed feedback episode that could motivate her to pay attention to it – for example, because it only contained information she was very familiar with already – she grew impatient with her peers’ errors.

Example 48: “I wish Zoe would say the right thing, because he keeps repeating the same stuff over and over, trying to get her to say what he wants her to say correctly. And I was thinking, oh my God, hurry up, so we can move forward!”
Global theme “Message”

In a total of fourteen comments, ten interviewees mentioned the meaning content of the feedback episode as taking precedence over its linguistic aspects in attracting their attention. Figure X illustrates this theme.

“What this learner was saying was very interesting for me.”

Interviewees felt more motivated to pay attention to the feedback episode if it contained genuine information about their classmates that they had not known.

Example 49: “I remember this very well cause I didn't know that he spoke Spanish and was from Texas.”
“What this learner was saying was very confusing to me.”

The content of a peer’s message could also be difficult to interpret, and thus attract and keep their attention.

Example 50: “I understood the whole sentence really good. So I was thinking:

“Why would somebody's parents like the hot weather?”

---

Figure 20. Thematic network “External distractors”.

Global theme “External distractors”

Finally, a number of comments mentioned as reasons for not paying attention to the classroom interaction because of factors unrelated to the corrective feedback episode itself.

“I was tired.”

A frequently mentioned – and intuitively logical – reason for failure to pay attention to what was going on in class, was fatigue. Interviewees recalled being tired most often while
commenting on the final few corrective feedback episodes in their sequence which had occurred
during the final third of the class period

Example 51: “I don't remember. It was towards the end of class and I was kind of
tired and kind of spaced out to a degree.”

“I was thinking about something else entirely.”

Trivial concerns about getting their study materials or homework ready were also
reported as reasons not to pay attention to the classroom interaction.

Example 52: “I was looking for a piece of paper to copy the exercise that I'd done
for the night and I don't remember him, really. I knew he was talking
but I wasn't paying enough attention to whatever he was talking
about”.

“I'd answered a question already, so I wasn't listening/focusing.”

Interviewees also reported that they perceived having already participated in the
classroom activity as a reason to relax and turn their attention away from the interaction.

Example 53: "At this point I was just happy that my turn is over, and I was just
listening to what other people were saying cause I no longer had to
think what I was gonna say."
Combining quantitative and qualitative data

The five global themes and the subordinate organizing themes were compared against the results from the quantitative analysis, in order to identify patterns of interconnectedness and potential explanations for increased or decreased awareness of corrective feedback. The following patterns emerged from this comparison:

The global theme “External distractors” theme was associated with lack of awareness of the corrective feedback.

Focusing on the content of a peer’s contribution, rather than on its linguistic form, described by the global theme “Message”, was associated with decreased awareness of correction and target.

The relationship between the global theme “Linguistic information” and awareness was complex. Some of the subordinate themes, particularly the ones related to the learner’s own experience with the same error or linguistic information were associated with increased awareness. However, the organizing theme “Thinking how to answer the question” was associated with lack of awareness of the corrective feedback occurring at the time, or perhaps lack of motivation to mention that feedback in the recall interview since it was perceived to be of lesser importance.

“Learner” was a global theme that was strongly associated with awareness or lack thereof. The belief that the participating learner had a high proficiency or abilities in Arabic was associated with an increase in reported awareness both of the occurrence of corrective feedback, and of its linguistic targets. In contrast, a reported dislike of the participating student, whether
related to their personality, performance, or undisclosed reasons, was associated with absence of awareness of the corrective feedback.

Finally, the global theme “Feedback Episode” was associated only with increased awareness, regardless of whether the learner reported positive or negative attitudes towards the teacher's error-correction method, or the specific instance of teacher-provided feedback. This somewhat counter-intuitive finding may be related to the fact that all learners who reported negative attitudes towards the teacher’s intervention were advanced learners. The quantitative analysis of the data shows that advanced learners are generally more likely to report awareness of both a correction, and its target, than beginners. Therefore, it is difficult to determine the exact relationship between this type of comments, and reported awareness. On the one hand, a possible implication of this finding is that any reason to focus on the teacher – both a positive evaluation of their behavior, and a negative one - increases the learner’s engagement with the classroom interaction and, consequently, awareness of the corrections provided by the teacher. On the other hand, it is just as plausible to suggest that advanced learners are, by virtue of their relatively larger experience and expertise, more cognizant of the interactional processing occurring in the classroom, to the extent that even negative attitudes to the teacher’s corrective intervention are not sufficient to occupy all their attentional resources and prevent them from becoming aware of a correction.
CHAPTER 5: DISCUSSION AND CONCLUSIONS

Introduction

This chapter will discuss the study’s findings in relation to the research questions and hypotheses. The results from the quantitative and qualitative analyses will first be summarized and then discussed in regard to relevant theoretical and empirical research. My analysis will show that cognitive factors – feedback characteristics such as explicitness and linguistic focus – are not a sufficient explanation for the differing rates of learner awareness of classroom feedback. Socio-affective factors, related to the learners’ perceptions of themselves, their peers, and the interactional context, will be shown to play an important role as well. The discussion of the findings will be followed by an overview of their theoretical, methodological, and pedagogical implications. The chapter concludes with a discussion of the study’s limitations and suggestions for future research.

Summary of findings

This study sought to investigate beginning and advanced Arabic learners’ awareness of classroom corrective feedback and the factors associated with differing rates of awareness. Specifically, the study addressed the relationship between awareness and the type, linguistic target, and direction of the feedback. A qualitative analysis of the data was also performed with the aim of identifying additional factors that may be related to learner awareness of feedback.
Quantitative analysis of the data showed that advanced learners reported awareness of correction and its target significantly more often than beginners. Both groups showed moderately high rates of awareness of correction – 75% for advanced, and 50% for beginning learners.

None of the factors under investigation – type, linguistic target and direction of the feedback – were found to influence advanced learners’ awareness of correction. For beginning learners, the only factor that significantly impacted awareness was the direction of the feedback – the learners were 30% more likely to report awareness of correction when the feedback episode had been directed at them.

The characteristics of the feedback had no impact on the advanced learners’ awareness of the linguistic target of the correction either. For beginning learners, however, both type and target of the feedback were significantly associated with differences in reported rates of awareness. Although the strength of this association (measured with Cramer’s V) had relatively low values - 0.19 for both - these numbers were interpreted as having high practical significance because they helped narrow down the vast number of factors with potential influence on awareness. Beginning learners were more likely to report being aware of the linguistic target of the feedback if the feedback had been directed at them, if it had been delivered in the form of explicit or combination feedback, and if it had targeted lexis. Declarative recasts, as well as feedback targeting phonology, were significantly less likely to be noticed.

The direction of the feedback influenced both advanced and beginning learners’ ability to notice the linguistic target of the corrective feedback. The learners from both groups reported awareness more often when the feedback had been directed at them.

In addition to the quantitative analysis, a qualitative thematic analysis was also applied in order to identify additional factors, related to the learners’ awareness of feedback. Five major
themes were identified in the learners’ stimulated recall comments: attitudes to the feedback episode, attitudes to other learners, linguistic information contained in the episode, the message that the learner in the episode was trying to convey, and, finally, external distracting factors. A notable finding from the qualitative analysis is that in most cases when the learners were not attending to the correction or the target of the episode, they were focusing on a different, arguably equally legitimate, feature of the feedback episode, rather than being distracted by factors unrelated to the classroom discourse.

**Proficiency as a factor in learner engagement with feedback**

The quantitative analysis showed significant differences between beginning and advanced learners in regard to their understanding of feedback. Advanced learners were much more likely to be aware of a correction, both when they were its direct recipient, and when it was directed at someone else in the class. This finding lends indirect support to Ammar & Spada’s (2006) claim that proficiency is an important factor in determining how language learners use the feedback that is presented to them. Two explanations for such difference can be advanced: expertise, and developmental stage.

The advanced learners in this study had received much longer exposure to both the linguistic content of the observed classroom interactions, and to corrective feedback related to linguistic errors in Arabic. It is reasonable to assume that as a result of this prolonged exposure, they had developed relative automaticity in processing these features of the interaction, and consequently had sufficient attentional resources “freed up” to satisfy higher processing demands. Beginning learners, on the other hand, may have been too preoccupied with the basic processing of linguistic information – comprehending and constructing phrases in Arabic. Thus,
they may not have had the necessary resources to focus on corrections and their linguistic targets, particularly when the corrections were addressed to someone else in the classroom.

The qualitative analysis of the learners’ comments highlighted another difference between the two groups: in attitudes. During the coding for the thematic analysis, several emotionally-charged comments from the advanced class stood out, in which the interviewee reported feeling exposed, targeted, and embarrassed by the teacher’s feedback. This was an unexpected finding, for the following reason: All participants, as part of completing their background questionnaire (see Appendix B), had answered questions about their attitudes towards receiving corrective feedback during classroom speaking activities. The answers were essentially identical across classes and proficiency levels: All learners felt that feedback was important, and that they personally benefited from the teacher’s feedback in their Arabic language class. Therefore, the negative responses were in stark contrast to the overwhelmingly positive reported attitudes.

A focused analysis of the negative comments, and subsequent cross-comparisons with feedback features and reports of awareness, showed that the negative reactions were reported by two learners in the advanced class and concerned both direct and indirect episodes. Example 54 shows such comments, preceded by the classroom episode to which they refer. Bolded phrases reflect participant emphasis.

Example 54: Negative reactions to corrective feedback, advanced classroom.

Zoe:  
*li-l... li-z-zawja, li-l-bint, li-'anna *qult al-Hamaa 'in hiya fii barnaamaj majester wa ashiya' kadhalik*

to the wife, to the girl, because *I told the mother-in-law that she was in a master's program, and such

T4:  
*li-'anna-haa qaalat li-l-Hamaa 'anna-haa fii barnaamaj majester. li*
ʾanna-haa qaalat, laysa li-ʾanna-haa qult. li-ʾanna-haa qaalat

Because she said to the mother-in-law that she was in a master's program. Because she said, not because I said, because she said

Zoe: qaalat

She said

T4: marra ʿuxraa, marra ʿuxraa, li-ʾanna-haa qaalat

One more time, one more time – because she said

Zoe: li-anna-haa qaalat inna al-bint hiya fii barnaamaj majester...

Because she said that the girl was in a master's program.

Stimulated recall comments:

Zoe (recipient of feedback):
I was again a little annoyed. I realized that after I said it, I realized that I had, you know, conjugated the verb incorrectly. You know, it should’ve been qaalat not qult, but I didn’t think that it made a huge difference and that people still understood. I was talking about the mother-in-law, I was talking about what she thought and what she was saying, and you know, I was a little embarrassed that he made me repeat it front of the class - felt like he was almost sort of rubbing in my mistake

Sean (observer):
At that time I was thinking.. why did Zoe say qult? I knew exactly what mistake he was gonna correct there because he just loves to correct grammatical mistakes that we make so I knew exactly what he was gonna say cause I knew the mistake she made, so.. and I was thinking God, I don’t wanna be here... 

It is interesting to note in this example that the corrective feedback episode that triggered the learners’ reactions was a combination episode – that is, it contained multiple turns and types of feedback, targeting a single error (in this case, the learner had conjugated the verb, qaala [to say] in the first person singular, rather than third person singular feminine form). An examination
of all episodes associated with this type of negative comments, showed that all of them (n=6) were combination episodes.

The two learners similarly reported negative attitudes to the teacher’s correction not only when the feedback was directed to them, but also when it was directed to another learner in the classroom. Example 55 presents such a comment:

Example 55. Negative reaction to feedback directed at others, advanced classroom.

Zoe: I was a little annoyed at professor T4 because Mary was trying to talk, trying to get something across - and he sort of stopped her in the middle for a grammar lesson. And I felt that that sort of disrupted the flow of what she was trying to say.. and .I’m sitting next to Mary and I - in the video I noticed I just kinda drop my head cause I was just like “oh God!” - you know, why did he have to do that?! You know.. she was getting somewhere we could understand her, why did he have to stop?!

It is evident from these examples that some of the advanced learners expressed strong negative attitudes towards the feedback they or their peers were receiving, when this feedback was in the form of a combination episode. It may be the case that this type of feedback is not suitable for more advanced learners because of its repetitive nature. The data from this study show that advanced learners are aware of both the correction and its target most of the time, so they may construe repetitive treatments of the same error as placing an unnecessary emphasis on the mistake – “rubbing it in”, as Zoe put it. In other words, the reported negative attitudes may be a reaction to one particular type of feedback, rather than to feedback as a whole – an explanation that reconciles the negative comments with the overall positive attitudes to feedback, reported through the background questionnaire.

In contrast, stimulated recall reports from beginning learners revealed only positive or neutral attitudes to corrective feedback, including to repetitive feedback of the combination type,
which occurred frequently in those classrooms. Example 56 shows a typical set of reports from the lower-proficiency learners:

Example 56. Reactions to feedback, beginning classroom.

Cathy:  

*bi-...*nisbatu-ka...maadhaa madiinat ’aHsan?

*In your opinion what city is the best?

Teacher:  

*bi-n-nisba laki

In your opinion

Cathy:  

*bi-n-nisbat laki

In your opinion

Cathy (recipient of feedback):

I had just asked Clare to remind me how you say ‘in your opinion’ and I want to say ‘in my opinion’, *bi-nisbat lii* So I thought I’ll just add the -uka or -uki at the end like a possessive. So I said *binisbatuka* and she corrected me. But I’m actually proud of that one because I thought I knew how to compose that question and it was *bi-nisbat laki* stuff I don’t know as well when I’m referring to someone else but yeah I felt that was good effort on my part.

Anne:  

Um I was prolly just listening to Cathy say stuff about her family. It’s really hard cause we don’t have a lot of vocabulary about that. And then of course after Cathy said stuff wrong, [the teacher] corrected her and then Cathy was like “oh yeah, it’s this way”. Like the same way all of us do, when we say something wrong and then she [teacher] says it right and then we’ll be like “Oh, now I get it” and then we’ll say it right. Yep.

Clare:  

Um I was just- I wasn’t sure – I thought it was interesting how our professor constructed the sentence in the beginning and how she used *bi-n- nisha laki*, which means “in your opinion” and I remember repeating that over and over in my head so that I would know how to say it next time.

Brian:  

Um, I guess Cathy was looking for the right words there and struggling and gasping. Once again [name of teacher] was on the ball – professor, I mean I would never call her [by name] to her face. *Ustaadha* [professor] was on the ball and she helped Cathy do that. I do actually remember that. I was just thinking “what is she trying to say?” and then sorta understanding, but then *ustaadha*
[professor] was like “bla-bla-bla” and then I was like, “oh okay, right there” Yeah, it works well . . .

Stimulated recall comments from the beginning learners reveal a different attitude to feedback: They see it as helpful, compassionate, reassuring, and indispensable. Van den Branden (1997) proposed that for learners of lower proficiency, corrective feedback may indeed be more necessary, because they, by virtue of their limited knowledge, may be more dependent on their interlocutor “to turn feedback into interactionally modified output” (p. 625). For advanced learners, on the other hand, more knowledge translates into higher self-reliance, and the teacher’s feedback is no longer seen as necessary or helpful.

This finding has interesting pedagogical implications. In surveys of attitudes towards corrective feedback, teachers often report being reluctant to give feedback, particularly to lower-level learners, for fear of triggering a negative affective response. The data from this study, however, suggest that while this may be a valid concern at the advanced levels of proficiency, for beginning learners feedback plays a positive and reassuring role, and teachers should not withhold this type of assistance from their low-proficiency learners.

*Type and target of the feedback*

The quantitative analysis of the learners’ recall protocols regarding the relationship between the type of feedback and rate of awareness, yielded somewhat surprising results. At both the beginning and the advanced levels, implicit types of feedback, such as negotiation for the beginners and direct recast for the advanced learners, were perceived more accurately than explicit feedback. This finding lends support to Lyster & Mori’s (2006) Counterbalance hypothesis, according to which types of feedback that counterbalance the overall instructional
focus of the classroom, are more perceptually salient and therefore more readily noticed by the learners. In the current study, the more implicit conversational types of feedback like negotiation and recast were more noticeable because they contrasted with the overall form-oriented focus of the classrooms.

*Direction of the feedback*

Nabei & Swain (2002) in a case study of one advanced EFL learner’s perception of classroom corrective feedback, reported that their learner, Keiko, did not pay attention to feedback that was directed at other learners in the classroom, because she was not interested in what they had to say. In contrast, Ohta (2000) reported that her beginning and intermediate-level learners of Japanese were actively engaged in the corrective feedback episodes that involved their peers, often quietly repeating the teacher’s feedback or the other learner’s response. In the present study, the observed learner behavior at both levels of proficiency was consistent with Ohta’s (2000) findings. One way to account for these differences, is to look at the pedagogical foci of the three classrooms (Lyster & Mori, 2006). Both Ohta’s study and the current one were conducted in form-focused classrooms, whereas Nabei & Swain’s observations took place in a meaning-focused conversation class for advanced college learners of English. The increased emphasis on grammar in the Arabic and Japanese classrooms may have pushed the learners to use every opportunity to focus on and practice language structures by repetition, including repetition of other learners’ utterances.

Not all other-directed feedback episodes were viewed in the same light and with the same amount of interest, however. Comments in the “Learner” thematic network from the qualitative analysis showed both positive and negative attitudes to other learners and their contributions.
Several types of reports could be identified in the data. First, a few learners reported paying attention to the contributions of someone else, because this learner was “good”. Previous research has shown that learners tend to view more positively and to incorporate more readily the linguistic knowledge they receive from peers whom they consider more proficient or knowledgeable than themselves (Kim & McDonough, 2009; Watanabe & Swain, 2008). In fact, Watanabe & Swain in a study of intermediate EFL learners engaged in cooperative writing found that it was not the learners’ objectively measured proficiency but their perceived proficiency that determined how much authority their contributions would command in the discussion. They concluded that “…how learners perceive each other’s proficiency difference and how they interact with their partner based on their perceived proficiency might have a greater impact [than measured proficiency]” (Watanabe & Swain, 2008: 12-13, emphasis in the original).

Learners in the present study also reported that interpersonal dynamics influenced the extent to which they attended to the feedback directed at other learners, and to their contributions. Strong negative attitudes to the other learner resulted in deliberate non-attention to the whole feedback episode, regardless of its type or target. Positive attitudes and reported emotional investment in the episode – e.g., when the other learner was a friend or a hard worker, were associated with increased attention to the corrective feedback episode. This finding highlights the often neglected role of interpersonal dynamics in classroom learning. Ehrman & Dornyei (1998), after extensive analyses of interactions between and among learners and teachers in the classroom and their effect on learning gains, suggest that interpersonal dynamics do not simply flavor learners’ classroom experiences, but in fact determine the learning outcomes: “learning is enhanced by good interpersonal relations or seriously hindered by dysfunctional interactions between teachers and students and among student… The learning
experience of every student and the effectiveness of every teacher is influenced by what goes among and between the people who populate the classroom.” (p. 4-5). More research is needed to understand how and to what extent the personal affective undercurrents constantly present in the classroom influence language learning, and how this influence can be controlled and directed for the benefit of the learners.

In sum, this study found that a number of affective factors related to the learners’ perceptions of the feedback or of the participants in a feedback episode, influenced the extent to which learners reported being aware of the corrective feedback. Previous studies on learners’ awareness of classroom feedback (Kim & Han, 2007; Mackey et al., 2007; Nabei & Swain, 2002; Ohta, 2000; Roberts, 1995) have largely assumed that only characteristics of the feedback and the context determine the extent to which learners notice feedback. This study highlights the importance of socio-affective factors, such as interpersonal dynamics, attitudes towards the feedback episode or the teacher’s choice of corrective technique, attitudes towards other learners, the learner’s own perception of own proficiency relative to peers, etc. All of these factors emerged in the learners’ comments in relation to their reported rate of awareness of corrective feedback, suggesting the need for future research to look beyond the cognitive paradigm when exploring learner-internal attentional processes.

Implications

This study has implications for second language acquisition theory and research methodology, as well as for second language pedagogy.
Theoretical and methodological implications

Findings from this study suggest that proficiency plays a major role in learner awareness of classroom corrective feedback. These findings lend indirect support to claims in developmental studies that differing learning results are related to different abilities to notice feedback (e.g. Ammar & Spada, 2006). From the point of view of SLA research methodology, these findings also highlight the need to take into consideration participant proficiency in studies examining the relationship between awareness of feedback and language learning. Typically, laboratory studies have not distinguished between learners across the intermediate spectrum of language proficiency, lumping together both lower-intermediate and higher-intermediate learners (see, for example, Loewen, 2005; Loewen & Philp. 2006). Findings from this study, however, suggest that combining of participant groups at different levels of language development may obscure important patterns in the data. Thus, it is advisable to take into consideration the possible effect of proficiency as an intervening variable or otherwise to limit participant selection to a single level of proficiency.

Pedagogical implications

On the basis of this study’s findings, several suggestions can be made regarding the provision of corrective feedback during classroom interaction.

First of all, it seems that advanced and beginning learners not only understand classroom feedback at different rates, but also have different affective reactions to the feedback they or their peers receive during classroom interaction. The beginning learners in the study reported either neutral or positive attitudes to classroom feedback, whereas comments from some of the advanced suggested that they perceived the feedback as an unnecessary interruption or
intervention. Learners at the advanced level reported particularly negative attitudes in response to combination episodes - feedback episodes which involved multiple turns but focused on a single linguistic target. To the learners, this repetitive treatment seemed redundant and excessive.

Beginning learners, on the other hand, tended to view corrective feedback as a helpful and welcome intervention on the part or the teacher, regardless of the length of the feedback episode. In contrast to the advanced learners, beginning learners often reported being reassured by the teacher’s feedback.

The above findings suggest that learners at different proficiency levels may exhibit differing degrees of acceptance towards corrective feedback, as well as different affective responses to the teacher’s intervention. Teachers at the advanced level should use combination feedback with caution. Findings from this study show that in the case of advanced learners, lack of response or failure to correct a non-targetlike utterance is not necessarily an indication of insufficient linguistic knowledge. Excessive repetitive feedback can feed negative attitudes towards teacher’s corrections, which in turn may influence the learners’ motivation, self-esteem, level of comfort in the classroom, and willingness to communicate – and ultimately impact the learner’s language development.

At the same time, teachers should not shy away from providing corrective feedback in beginning classrooms for fear of hindering the learners’ emerging speaking abilities and confidence. Findings from this study suggest that feedback at this level is generally well-received and viewed as essential, helpful, and reassuring.

Participant comments also highlight the influence of factors such as task set-up and task timing on the learners’ ability to attend to corrective feedback. Specifically, learners showed a tendency to pay more attention to feedback if they were anticipating to be called on to
participate. When learners sensed an order to the way the teacher nominated them for participation, they reported “tuning out” after completing their turn, because they felt safe they would not be called on again. Teachers may be well advised to organize communicative activities in such a way that the learners’ participation would not follow a predictable pattern, so as to avoid this creating a sense of complacency among the learners, which could lead to decrease in attention.

Learner comments also showed that learners were more likely to “tune out” and stop paying attention to feedback during the last third of the class period. Hence, it may be best to plan feedback-heavy conversational activities with considerable demand on attentional resources in such a way that they do not fall too close to the final third of the class period.

Limitations of the study

This study has several important limitations which need to be taken into consideration when interpreting the results. First, the study was conducted in language classrooms in a large Arabic language program at a highly selective US university. Hence, the learners’ motivation and the teachers’ pedagogical approach may not be representative of Arabic teaching contexts in the US as a whole. The classroom context for the study was largely form-focused, rather than meaning focused, and in this respect the study differs in significant ways from the typically communicative, conversation-based, fluency-oriented language classrooms in which much of the research on corrective feedback perceptions and efficacy has been conducted. Thus, findings from this study cannot be generalized to include learners from more communicatively oriented classrooms. Neither can they be generalized to include younger learners or learners of languages other than Modern Standard Arabic.
Second, the sample sizes in this study were small (n total = 31, n beginning = 26, n advanced = 5), particularly in the case of the advanced learners group, due to logistical constraints of the data collection. Because the program in which data collection took place hosted significantly fewer classes at the advanced level, than at the beginning level, recruiting comparable sample groups at both levels turned out to be challenging. Since the implementation of the stimulated recall measure depended on voluntary participation of the learners, participant attrition in the study was very high, and the measures taken (e.g., accommodating individual schedules and offering gift certificates as compensation for participation) were insufficient to offset this trend. Larger sample sizes would increase the robustness of the quantitative results by minimizing the effects of individual differences.

The same logistical constraints applied to the selection of classrooms for participation in the study. While at the beginning level data could be collected from four intact classes with three different instructors, at the advanced level the data collection took place in a single classroom, with only one instructor. Individual differences pertaining both to the learners and to the instructors could thus significantly influence the findings from this study.

The small sample did not allow for the controlling of two important variables in the selection of the classroom instructors who participated in the study: gender and native speaker status. The three instructors at the beginning level were all female, and the advanced level instructor was male. Gender can play a role in the instructor’s general attitude to feedback, as well as to the specific classroom climate, with male instructors typically providing less feedback and fewer opportunities for learners to respond to it (Fassinger, 1998). Research on the provision of feedback by native-speakers and non-native speaker language teachers has also identified divergent patterns. Non-native speakers are more prone to provide abundant feedback,
particularly in explicit form, while native speakers show a preference for implicit feedback, particularly recasts (Medgyes, 2001). These differences may have potentially influenced the results from the study, and should be taken into consideration in future research.

Third, it is also important to acknowledge the limitations of the study related to the data collection methodology and the nature of stimulated recall reports. As discussed in previous chapters, verbal reports equate reported information with perceived information, i.e., in the use of verbal reports, the researcher assumes that learners will report all their thought processes and only their thought processes. However, as discussed repeatedly in research employing such measures of noticing, the absence of evidence is no evidence of absence. In other words, the fact that a particular phenomenon or incident were not mentioned in a learner’s report does not necessarily indicate that they were not noticed by the learner. For this reason, lack of awareness is particularly difficult to assume in studies using recall data.

Finally, another important limitation is related to the statistical tests used in the study. The small sample sizes and the nature of the data necessitated the use of a non-parametric statistical test, chi-square. Non-parametric tests are widely used in language research, particularly in classroom research which typically involves small participant numbers. However, these tests, while more versatile and accommodating than parametric statistics, also have lower predictive value. A larger sample size, combined with parametric statistical tests, would increase the robustness of the findings.
Suggestions for future research

Areas for improvement in the current study include a larger sample size, a more heterogenous participant base, and controlling for intervening variables related to the teachers. It is also important to point out that this was not a developmental study, that is, it makes no claims regarding the relationship between awareness of corrective feedback and language development. Future research should address empirically the relationship between contextual and learner-specific factors, awareness of feedback, and learning outcomes in a unified design. Another important direction for future exploration is related to the cross-sectional nature of this study, in which the investigation of awareness of feedback was limited to a single class. If learner perceptions and attitudes do, as some SLA research suggests, change over time, this may be a worthwhile question to pursue in future longitudinal research.

Conclusions

The purpose of this study was to investigate the interplay between beginning and advanced learners’ awareness of corrective feedback during classroom interaction on the one hand, and the linguistic, typological, and contextual characteristics of the feedback (type, target, and direction of the feedback), on the other.

To summarize, proficiency was shown to play a major role in learner awareness of feedback. Advanced learners reported awareness of both correction and linguistic target of the feedback at a significantly higher rate than beginning learners.

Furthermore, advanced learners’ awareness of corrective feedback was not constrained by any of the feedback-specific variables investigate in the study – type, target, and direction. It seems that at the level of language development at which these learners were, interaction may
have a much more limited effect on learning than it does at the beginning levels. Garcia-Mayo and Pica (2001) voice a similar concern – if interaction, particularly communicative breakdowns that trigger output modifications, is facilitative for learning, then the fact that both interaction and communicative teaching and learning come up insufficient when it comes to providing advanced learners with opportunities for learning, is of some concern for language teachers and practitioners alike. Moreover, this finding may suggest a ceiling effect for higher proficiency on the developmental benefits of interaction, consequently questioning the explanatory power of the interaction approach for the learning that takes place at the advanced levels of proficiency and beyond.

Another set of important findings in this study comes from the qualitative analysis of the data. Research on learner awareness of feedback has so far placed an emphasis on the importance of paying attention to feedback. In a classroom, however, a lot is going on at the same time and there are competing demands for the learner’s attention. The current study shows that while the learners in the study missed many opportunities to notice feedback, often this was due to parallel occupation with other facets of classroom work, arguably equally important – new linguistic information, for example. They did not always see the feedback in the corrective episode as the single most important focus for their attention. Frequently, the reason they were not thinking about the feedback they were witnessing was that they were focusing on other linguistic, informational or semantic aspects of the episodes – e.g., thinking about lexical items that were new or other new linguistic information, confirming their hypotheses not coinciding with the target of the feedback. To sum up, findings from the qualitative analysis show that lack of awareness of feedback is not necessarily associated with lack of attention to the linguistic information in the episode itself – the learners simply looked at it from a different angle.
This study highlights the need for further research in two main directions: the socio-affective factors that impact learner engagement with corrective feedback, and the effect of proficiency on attention and awareness. With the increasing focus on the development of advanced language capabilities as a response to changing geo-political currents, such research is particularly relevant at the present moment.
## Appendix A

### Transcription and Glosses

#### Transcription Chart

<table>
<thead>
<tr>
<th>Arabic script</th>
<th>IPA</th>
<th>Phonemic transliteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>ء</td>
<td>?</td>
<td>'</td>
</tr>
<tr>
<td>ب</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>ت</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>ث</td>
<td>θ</td>
<td>th</td>
</tr>
<tr>
<td>ج</td>
<td>ð</td>
<td>j</td>
</tr>
<tr>
<td>ح</td>
<td>h</td>
<td>H</td>
</tr>
<tr>
<td>خ</td>
<td>χ</td>
<td>x</td>
</tr>
<tr>
<td>د</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>ذ</td>
<td>ð</td>
<td>dh</td>
</tr>
<tr>
<td>ر</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>ز</td>
<td>z</td>
<td>z</td>
</tr>
<tr>
<td>س</td>
<td>s</td>
<td>s</td>
</tr>
<tr>
<td>ص</td>
<td>š</td>
<td>S</td>
</tr>
<tr>
<td>ض</td>
<td>ð</td>
<td>D</td>
</tr>
<tr>
<td>ط</td>
<td>ɹ</td>
<td>T</td>
</tr>
<tr>
<td>ظ</td>
<td>Ɂ</td>
<td>'</td>
</tr>
<tr>
<td>ع</td>
<td>ŋ</td>
<td>gh</td>
</tr>
<tr>
<td>ف</td>
<td>f</td>
<td>f</td>
</tr>
<tr>
<td>ق</td>
<td>q</td>
<td>q</td>
</tr>
<tr>
<td>ك</td>
<td>k</td>
<td>k</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arabic script</th>
<th>IPA</th>
<th>Phonemic transliteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>ل</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td>م</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>ن</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>هـ</td>
<td>h</td>
<td>h</td>
</tr>
<tr>
<td>و</td>
<td>w</td>
<td>w</td>
</tr>
<tr>
<td>ي</td>
<td>j</td>
<td>y</td>
</tr>
</tbody>
</table>

#### Vowels and diphthongs

<table>
<thead>
<tr>
<th></th>
<th>Arabic</th>
<th>IPA</th>
<th>Phonemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>a:</td>
<td>ːaː</td>
<td>aa</td>
<td></td>
</tr>
<tr>
<td>u:</td>
<td>ːuː</td>
<td>uu</td>
<td></td>
</tr>
<tr>
<td>i:</td>
<td>ːiː</td>
<td>ii</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>ːaː</td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>u</td>
<td>ːuː</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>ːɪː</td>
<td>i</td>
<td></td>
</tr>
<tr>
<td>aw</td>
<td>ːawː</td>
<td>aw</td>
<td></td>
</tr>
<tr>
<td>ar</td>
<td>ːarː</td>
<td>ay</td>
<td></td>
</tr>
<tr>
<td>wa</td>
<td>ːwaː</td>
<td>wa</td>
<td></td>
</tr>
<tr>
<td>ra</td>
<td>ːraː</td>
<td>ya</td>
<td></td>
</tr>
</tbody>
</table>

#### Abbreviations and symbols

- *: non-targetlike segment
- 3: third person
- fem: feminine
- f-II: verb form (derivation pattern) II
- fut: future
- gen: genitive
- imper: imperative
- imperf: imperfective
- indic: indicative
- masc: masculine
- neg: negative particle
- perf: perfective
- pl: plural
- rel: relative pronoun
- ques: interrogative particle
- sing: singular
- subj: subjunctive
Notes on transcription and transliteration:

All Arabic examples and quotes appear in italics. English translation for Arabic words that appear in interview comments is given in square brackets after the target word, e.g.: “I did not know the word *Hamaa [mother-in-law]*”. In examples of classroom interaction, a gloss with idiomatic English translation is provided underneath the transliterated Arabic text. Whenever the example necessitates that the morphosyntactic structure of a word or sentence be indicated, a second gloss-line is provided between the Arabic text and the English translation.

Example:  

*hunaaka waaHid shay laazim nunaaqish*

(there is *thing-one necessary we-discuss)  

(Morphosyntactic gloss)

there is *one-thing we should discuss  

(English translation)

Non-targetlike forms identified by the teacher as the focus of her feedback are marked with an asterisk (*). To avoid extraneous detail, non-targetlike forms in teacher utterances, as well as learner errors that were not addressed in the feedback, are not marked in any way.
Appendix B
Consent Form

Georgetown University Social & Behavioral IRB
IRB Number: 2006-513
Title of project: Learners’ perceptions about corrective feedback in the Arabic second language classroom

GEORGETOWN UNIVERSITY CONSENT TO PARTICIPATE IN RESEARCH

PROJECT NAME
Learners’ perceptions about corrective feedback in the Arabic second language classroom

PROJECT DIRECTOR
Gergana Atanassova

PRINCIPAL INVESTIGATOR
Gergana Atanassova

TELEPHONE
202-460-9381

SPONSOR
N/A

The Georgetown University Institutional Review Board (IRB) has approved this research project. For information on your rights as a research subject, call the Institutional Review Board office at 202-687-1506.

INTRODUCTION

You are invited to participate in a research study. I will be investigating teachers’ and learners’ perceptions in relation to Arabic language instruction. This form will describe the purpose and nature of the study, its possible risks and benefits, and your rights as a participant in the study. Please take whatever time you need to discuss the study. If you decide to participate, please sign and date the last line of this form.

BACKGROUND AND PURPOSE OF THE STUDY

The purpose of the study is to understand teachers’ and learners’ perceptions about classroom activities.

NUMBER OF PARTICIPANTS

Approximately 40 learners and teachers of Arabic will participate in this study.
PLAN OF THE STUDY

Participants will engage in normal classroom activities which will be video-taped. After the class, the participants and teachers will be asked to watch the video and make comments about what they were thinking during the original activities.

HOW TREATMENT WILL BE DETERMINED IN THIS STUDY

There is only one research group in this study. If you have any questions, concerns or in the event of an emergency, you may contact Geri Atanassova by telephone: (202) 460-9381 or by email: gsa4@georgetown.edu.

LENGTH OF STUDY FOR EACH PARTICIPANT

Learners and their regular classroom teacher will have a standard 65-min or 60-min Arabic class. Following class, learners and teachers will watch a video-recording of the same class and comment on it. The total duration of the study is expected to be up to 2.5 hrs.

POSSIBLE BENEFITS FOR PARTICIPANTS IN THE STUDY

Participants will receive gift certificates for the amount of $10 immediately following their interview. Participating learners will also have the opportunity to see a video-tape of the class, which may give them an extra opportunity to notice language that they missed the first time around.

POSSIBLE SIDE EFFECTS AND OTHER RISKS OF PARTICIPATING IN THE STUDY

There are no risks to participating in this study.

WHO CAN PARTICIPATE

This study is designed for learners of Arabic.

CONFIDENTIALITY OF THE DATA COLLECTED DURING THE STUDY

Wherever possible, each participant’s identity will be kept confidential in all reporting of this research, whether at conferences or in journals or book chapters, or other publications. After collection of the data, the researcher will immediately assign all your data a code and will never intentionally identify you. In addition, the identity of your language program will be kept confidential.

Page 2 of 3
THE RIGHTS OF PARTICIPANTS IN THIS STUDY

Participation in this study is entirely voluntary. You have the right to cancel your participation at any time. Leaving the study will not result in any penalty of any kind. Should you decide not to be included in the research, you are requested to notify the investigator about the cancellation. You can contact the investigator by telephone, through a letter or in person. There will be no consequences for you or anyone in your class or your program who decides to cancel participation.

PROBLEMS AND QUESTIONS

Call Geri Atanassova 202-460-9381 day or night if you have questions about the study. You may also call the Georgetown University IRB office at 202-687-5594 with any questions about your rights as a research participant.

INVESTIGATOR'S STATEMENT

I have fully explained this study to the participant. I have discussed the procedures and treatments, the possible benefits, the standard and research aspects of the study, and have answered all of the questions that the participant has asked.

Signature of investigator ___________________________ Date _________________

PARTICIPANT’S CONSENT

I have read the information provided in this Informed Consent Form. All my questions were answered to my satisfaction. I voluntarily agree to participate in this study.

Signature of witness ___________________________ Date _________________

Your signature ___________________________ Date _________________
Appendix C

Learner Questionnaire

1. Name/Code:
2. Gender:
3. Age:
4. Major:
5. First language(s):
6. How long have you studied Arabic?
7. How old were you when you started studying Arabic?
8. Why are you studying Arabic?
9. Have you studied other languages besides Arabic? For how long? What is your proficiency level in those languages?
10. Where have you studied Arabic (Mark all that apply): Home, school, university, religious language school, private tutoring, study abroad (country), other (specify).
11. What varieties of Arabic (MSA, dialect) have you studied so far? If you have studied a dialect, please specify which dialect:
12. Are you getting additional instruction in Arabic now? Where? How often?
13. Do you interact in Arabic outside of class? With whom? How often?
14. Do you think that speaking Arabic with others is helpful? If yes, how does it help? If no, why?
15. Do you think that getting feedback from your teacher while you are speaking in class is helpful? If yes, how does it help? If no, why?

16. The interview you just participated in was part of a research study. What do you think the purpose of the research study was? When did you discover that?

17. Any other comments?
Appendix D

Instructions for stimulated recall

Before the start of each stimulated recall interview, the following instructions were read to the learners and teachers who had volunteered to participate. Phrases given in bold were emphasized and/or repeated during the reading.

I am going to show you several video clips, taken from your Arabic class this morning. The clips will be very short, sometimes as short as 5 seconds, and every clip will contain a situation that happened during your class. Watch carefully and try to recall the moment when this situation took place. **What were you thinking when this was happening?**

After every clip, I will ask you to turn to your computer, speak loudly and clearly in English into the microphone and say what you were thinking when this situation took place in the classroom. Remember, I ask you to tell me what you were thinking **back then**, when this was happening in the classroom, **not** what you are thinking **now**.

If you cannot recall what you were thinking, just say “I don’t remember”. There are **no** right and wrong answers: Just say what you were thinking **even** if it had nothing to do with the class or with Arabic. For example, it is okay to say you were thinking that you were hungry, if this is really what you were thinking. It is also okay to say that you weren’t thinking anything, if that’s the case.

Your answer can be as long or as short as you want it to be. Do not worry if the others finish their comments before you: just continue talking until you have said everything you can recall.

Remember: Tell me **only** what you were thinking **while this was happening** in the classroom, **not** what you are thinking now. If you **do not remember** what you were thinking, say “**I don’t remember.**”

Do you have any questions?
REFERENCES


In K. M. Bailey & D. Nunan (Eds.), *Voices from the language classroom: Qualitative research in second language education* (pp. 168-196). New York: Cambridge University Press.


(Eds.), *Introspection in second language research* (pp. 24-53). Clevedon/Philadelphia: Multilingual matters.


call for momentary assessments. In A. Baum, T. A. Revenson & J. E. Singer (Eds.), Handbook of health psychology (pp. 405-413). Mahwah, NJ: Lawrence Erlbaum Associates.


Medgyes, P. (2001). When the teacher is a non-native speaker. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 429-442): Henle & Henle.


